BMJ Open

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-021961
Article Type:	Research
Date Submitted by the Author:	08-Feb-2018
Complete List of Authors:	Mistry, Punam; University of Birmingham, School of Pharmacy Stirling, Heather; University Hospitals Coventry and Warwickshire NHS Trust, Paediatrics Callens, Claire; NIHR Clinical Research Network:West Midlands – Young Person's Steering Group (YPSG), CRN: Children Hodson, James; Queen Elizabeth Hospital Birmingham, Institute of Translational Medicine Batchelor, Hannah; University of Birmingham, School of Pharmacy
Keywords:	THERAPEUTICS, PAEDIATRICS, medicines, palatability

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines

Punam Mistry[1], Heather Stirling[2], Claire Callens[3], James Hodson[4] and Hannah Batchelor[1] on behalf of SPaeDD-UK project

(Smart Paediatric Drug Development – UK, accelerating paediatric formulation development http://www.paediatricscienceuk.com)

[1] School of Pharmacy, Institute of Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, UK

[2] University Hospital Coventry and Warwickshire NHS Trust, Clifford Bridge Road, Coventry, CV2 2DX, UK

[3] NIHR Clinical Research Network West Midlands, Institute of Research and Development, Birmingham Research Park, Vincent Drive, Birmingham, B15 2SQ, UK

[4] Institute of Translational Medicine, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston, Birmingham, B15 2GW, UK

Affiliation: Pharmacy and Therapeutics, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT, United Kingdom

Corresponding author: Hannah Batchelor, School of Pharmacy, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT
Email: h.k.batchelor@bham.ac.uk
Phone: +44 (0)121 414 3717

Abstract

Objective: To evaluate the age-appropriateness and suitability of patient reported outcome measures to assess the acceptability of the taste of oral liquid medicines in children.

Design and setting: An observational mixed methods study involving children aged 2-16 years old taking oral liquid medicine in paediatric inpatient wards across the West Midlands (UK). Assessment tools included: patient-reported scores on the taste of medicines via a 5-point facial hedonic scale; a visual analogue scale (VAS); a question, "Did you think the medicine tasted OK?" and researcher observations of facial expressions and behaviours immediately before, during and after administration.

Results: 611 children participated. The percent unable to complete the scales were 7% (n=46) for the VAS; 2% (n=15) for the hedonic scale and 1% (n=7) for the question about taste. Significant correlations (Spearman's rho) were observed between the patient reported outcome measures: 0.80 and 0.78 for the taste question and hedonic and VAS respectively and 0.84 for the hedonic and VAS. Researcher observations demonstrated the ability of the patient to take the medicine as intended but did not provide sensitive measures of taste. 5% of administrations were not taken as intended by the children. Medicines known to have poor taste (clarithromycin and prednisolone) showed mean hedonic and VAS scores of ≥3.5 and >65mm respectively.

Conclusions: Patient reported outcome measures correlate with each other and are a useful means to assess the taste (and acceptability) of medicines. Hedonic scales are better understood by children and should be the first choice tool in the assessment of medicines taste.

Key words: medicines, taste, acceptability, age-appropriate, hedonic scale, paediatric medicine, patient-reported outcome measures, VAS

Strengths and limitations of this study

- This is the first study to compare methodologies to assess the acceptability of taste of liquid medicines in a large UK based paediatric population aged 2-16 years
- The sample size in this study provided large data sets of six key medicines, which provides a representative comparison for newly developed products
- This study was conducted within an inpatient environment and the acceptance of taste of medicines in a domiciliary environment may differ
- The study design captured the most relevant aspects of acceptability of taste whilst minimising the burden to participants, it was not possible to measure every aspect.

Suggestions for future research include measurement of: impact of the devices used to administer medicines; use of alternative hedonic scales (e.g. those with 2-9 faces); alternative anchor phrases; an endpoint of neutral and not positive within a hedonic scale; further exploration of medicines that tasted OK as well as those with a reported negative taste.

Article Summary:

What is already known about this subject?

- New medicines for children must be demonstrated to be acceptable to a paediatric population
- Measurement of acceptability of the taste of medicines is complex and a wide range of methods have been used previously, making comparison between studies complex.
- There is a need for age-appropriate reproducible and reliable methods to measure the acceptability of medicines

What this study adds

- Patient reported outcome measures offer a pragmatic means to assess the taste of medicines
- 5-point hedonic scales were better understood compared to visual analogue scales in children aged 2-16 years
- Although 41% of medicines were reported to have unacceptable taste only 5% were so bad that they could not be taken as intended

Funding Statement

This work was conducted as part of the SPaeDD-UK project (Smart Paediatric Drug Development – UK, a project co-funded by Innovate UK and the contributing companies of AstraZeneca, Bristol Myers Squibb, GlaxoSmithKline, Juniper Pharmaceuticals and Pfizer.

(http://www.paediatricscienceuk.com).

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Data sharing statement

Additional data is available in the Supplementary files. The full data set is held by the corresponding author, please email with any requests for extra data.

Authorship contributions

Punam Mistry contributed to the acquisition, analysis and interpretation of the data.

Heather Stirling contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Claire Callens contributed to the design of the study and acquisition of the data

James Hodson contributed to the design, statistical analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Hannah Batchelor contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments. She is the corresponding author for this work

Word count of manuscript body: 3836

1. Introduction

Measurement of acceptability of paediatric medicines is important for two major reasons: (i) poor acceptability is likely to be associated with poor compliance and therefore suboptimal therapy; (ii) regulatory guidelines dictate that paediatric medicines' acceptability should be assessed (preferably in children) [1]. Patient acceptability has been defined as, "an overall ability of the patient and caregiver (defined as 'user') to use a medicinal product as intended (or authorised)" [1]. A major barrier in the assessment of acceptability and subsequent interpretation of results is the lack of guidance on how an acceptability test should be conducted and the criteria used to determine acceptability in children [2]. It is unlikely that a single acceptability test will be valid for all medicines, due to the need to consider the benefit:risk balance on an individual basis.

Taste was previously identified as the biggest barrier in adherence and compliance in a paediatric population [3]. Previous techniques to evaluate acceptability focussed on taste and were based on methodology adapted from sensory analysis undertaken in the food industry. A variety of scales have been used to measure the acceptability of medicines, yet it remains unclear whether any of the scales is better for a particular purpose with regard to validity, reliability, feasibility, and preference [4, 5]. Patient reported measures of taste offer a pragmatic means to collect data on products. However, there is often concern about the reliability of these measures, particularly in a paediatric population [6]. Sensory analysis is commonly used to show a preference for one product over another, but this is not always relevant to medicines as often there is no alternative available. In addition, it can be rationalized that the therapeutic benefits of medicines outweigh the negative taste. It is complex to distinguish between acceptability and likeability of a medicine's taste, yet this difference may be critical for appropriate interpretation of resulting data.

The validity of any new method relies on how well it compares to other measures [7]. The development of a reliable method to assess the taste of medicines requires comparison of a variety of tools. This study compares a range of methods to assess the taste of medicines within a large

paediatric inpatient population. Independent researcher observations are also compared to the self-reported data. The results of this study will be used to propose a suitable method that can be used for future taste assessments.

2. Materials and Methods

Three patient-reported outcome (PRO) measures were compared to each other, and to researcher observations of medicines administration in an observational mixed methods study. Bespoke PRO tools were developed for this study based on previous methodologies and in consultation with the National Institute for Health Research (NIHR) Children Specialty's Young Person's Advisory Group (West Midlands) [8]. The young people (aged 11-18 years) reviewed the tools and provided feedback that the tools were age-appropriate.

The hedonic scale selected was a genderless image where the mouth was the only expressive facial feature (see Figure 1(a)). The images were yellow to ensure they stood out from the background paper. They were obtained from S-cool the revision website (http://www.s-cool.co.uk/gcse/food-technology/systems-and-control/revise-it/sensory-evaluation (accessed December 2015)). Children and young people preferred simple faces and felt that this would be most appropriate for the youngest age group [8]; anchor phrases were not included on the hedonic scale to keep this tool basic. The visual analogue scale (VAS) included anchor phrases selected by the young persons' group as the most clear and relevant [8], these were used at the extreme ends of the continuous scale.

The direction of change was from positive to negative, which corresponds to the extensive data on hedonic scales used for pain assessment in children [9]. Previous studies in adult studies of foods showed no difference based on structural variations that read from positive to negative or vice versa [10].

The third PRO was a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure.

2.1. Participants and Setting

Paediatric inpatients aged between 2 to 16 years taking any oral liquid medicine as part of their medical care were recruited from inpatient wards at 11 sites across the West Midlands. The study was conducted between December 2015- April 2016. Ethical approval was granted by London Surrey borders ethical committee (REC reference: 15/LO/1253; IRAS project ID: 179684).

Demographic information was obtained on participant's age and whether this was their first dose of the medicine under evaluation. The medicine name; brand; manufacturer; concentration (strength); dose administered and product batch number was recorded.

Each participant was observed by a researcher as they took their medicine. Some medicines were provided to the patient as an oral liquid following extemporaneous preparation within the clinical setting, for example, dispersion of prednisolone tablets immediately prior to dosing. Depending on the capability of the child, the medicines were either self-administered, or administered by nursing staff and/or parents. Participants were asked not to mix the medicine with any other food product, as this might influence the participant's responses.

2.2. Patient-reported outcome tools

Participants were asked to complete (a) 5-point facial hedonic scale and (b) 10cm VAS (Figure 1) immediately after administration of their medicine; both scales were provided on separate paper documents in a randomized order. Children were free to ask for support in completing the questionnaires from parents, nursing staff or the researcher present. Both reporting documents included a third PRO (Figure 1 (c)) as a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure. The purpose behind this question was to endorse the reliability of the participant's reporting from both scale-based questionnaires.

Figure 1

The data from the scales (a and b) were transformed into numbers for subsequent statistical analysis. Hedonic expressions were converted to scores of 1 to 5 with 1 being the most positive expression and 5 the most negative. VAS scores were reported as measurements from the extreme left hand side of the scale in mm; this provided scores of 0-100 mm (although scale was written as 0-10 cm).

2.3. Researcher Observations

A facial expression and behavioural scale to capture researcher observations was a 12-point tick chart designed in-house (Figure 2), based on an existing scale used to measure behavioural distress in children undergoing medical procedures [11], the Procedural Behavior Rating Scale (PBRS). The PBRS was developed specifically for infants and children with cancer, who were undergoing bone marrow aspirations or lumbar punctures. It contained 13 behavioural codes including: cry, cling, pain verbal, flail, stall, scream, refuse position, restrain, emotional support, muscular rigidity and requests termination [11]. The in-house tool was based on this scale plus other scales used to measure behavioural distress in children [12]. The facial expressions included on the scale were derived from previous studies that assessed food-liking in children based on their facial expressions; typically negative tastes are associated with eyes squeezing, brow bulge, nose wrinkle and pursed lips [13].

Figure 2

A series of nine short films or still pictures of children were made available to researchers participating in the study to assess the inter-rater agreement in the facial expressions displayed.

2.4. Statistical analysis

A sample size was not fixed for this study at the outset as there was no appropriate power calculation. Advice from a statistician was to ensure we had sufficient numbers (<50) paired data to ensure that selectivity and sensitivity of the methods could be demonstrated.

Initially, the mean age of participants that were unable to understand the assessments were compared to those who could using Mann-Whitney tests. The same approach was also used to compare the scores for patients receiving their first dose, relative to those who had previously received the medicine. Age was then divided into categories, and the proportions of participants scoring in the extreme categories for the scales were compared using Fisher's exact tests.

Spearman's correlation coefficients were used to assess the degree of correlation between the assessments for the cohort as a whole, and within each of the age categories. In this analysis, the "Did you think the medicine tasted OK?" question was treated as an ordinal scale, with the response of Don't Know being between Yes and No. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available

The assessments were then dichotomised, and compared using McNemar's tests to assess for marginal homogeneity in the 2x2 tables, and Kappa to assess agreement. The three assessments were then combined into a composite score, which was compared with reported behaviours using Kendall's tau correlation coefficients. All analyses were performed using IBM SPSS 22 (IBM Corp. Armonk, NY), and $p<0.05$ deemed to be indicative of statistical significance throughout.

3. Results and discussion

Data were available for 628 administrations to 611 children aged between 2-16 years. The median participant age was 6 years. Further details on the distribution of the participant ages can be found in supplementary material 1.

To ease analysis of data, the population for this study was stratified by age into three groups: 2-4 years (n=237); 5-9 years (n=227) and 10-16 years (n=147).

The medicine was administered as the first dose in 162 cases. There was no evidence of a significant difference in the hedonic or VAS scores between those receiving their first dose of a medicine, compared to those who had previous administrations (p=0.336, 0.909 respectively). For all subsequent analysis the data was pooled for those receiving their first and subsequent doses of medicine.

3.1. Completeness of patient-reported assessment scales

The assessment scales were not completed by all of the study participants. The VAS had the lowest completion rate, where 46 (7%) were not completed due to lack of understanding by the child, compared to 15 (2%) for the hedonic scale. There were 7 (1%) cases where the child did not understand the question, "Did you think the medicine taste OK?". The range and mean age of those unable to complete the scales were 2-5 years with a mean of 2.4 years for the hedonic; 2-6 years with a mean of 3.1 years for the VAS and 2-3 years with a mean of 2.3 years for the taste question. In each case, participants unable understand the assessment methods were significantly younger than the remainder of the cohort ($p<0.001$ each assessment). The cognitive function of children was not assessed within this study and there was an assumption of cognitive normal for age for all participants.

3.2. Distribution of responses to patient-reported assessment scales

In the patient reported scales (VAS and hedonic), the responses tended to be biased towards the extreme responses, with 56% of responses being in the highest or lowest categories for hedonic score (Figure 3a) and, 55% within 1cm from either end of the scale in the VAS (Figure 3b).

Figure 3

The use of the extreme ends of the scales was greater in the younger populations (p<0.001) (data is shown in Supplementary material 2). For those that completed the hedonic scale, 64% of those aged 2-4 years used the extreme ends of the scale compared to 60% of those aged 5-9; 35% of those aged 10-16 years. A similar trend was observed in the VAS data where 66% of those aged 2-4 years were within 10mm of either end of the scale compared to 60% of 5-9 year olds and 39% of 10-16 year olds.

3.3. Reliability of assessment, "Did you think the medicine tasted OK?"

The question, "Did you think the medicine tasted OK?" was asked to participants on two occasions for each medicine administration; once following the hedonic and once following the VAS (which were presented in a randomized order). In the 600 cases where participants responded to both questions, the same response was given in 569 (95%) of cases, giving a Kappa statistic of 0.91. In cases where the answers were not consistent, the median age of the respondents was 5 years (mean age = 6.3 years). For subsequent analysis, the two versions of this question were merged, with discrepancies resolved by taking the most positive response given by the participant.

3.4. Correlation between patient reported outcome measures

Significant correlations were observed between the hedonic scale score, VAS and "Did you think the medicines tasted OK?" question (all p<0.001 Table 1), with the strongest correlation observed between the hedonic and VAS scores (Spearman's rho=0.84). The weakest correlations were consistently observed in the youngest patients (age 2-4 years), implying that this group of patients had the lowest consistency in scores given across the different assessments. However, despite this, the consistency between the scores was still reasonable, with correlation coefficients ranging from 0.68- 0.77.

The assessments were dichotomised, to identify those responses that classified a medicine as tasting acceptable. For the hedonic scale, this was defined as a neutral to positive response (the leftmost three faces of Figure 1a). For the "*Did you think the medicine tasted OK?"* question, grouping the "yes" and "not sure" responses was considered most appropriate, as the overall purpose is the assessment of acceptability of taste of a medicine; the response "no" is clearly aligned to an unacceptable taste, making other responses acceptable. The level of agreement between the resulting dichotomised scores was high (90%), with a Kappa statistic of 0.782, and no indication of bias (McNemar's test: p=0.519).

Correlation of the hedonic scale to the "tastes OK" question is simple; however interpretation of the VAS is more complex. The VAS had anchor phrases at either end to maintain the continuous scale, so it was not obvious where neutral or positive was ranked on the scale. Initially, an arbitrary cut-off of <50mm in the VAS was used to define acceptability, which gave reasonable agreement with both the dichotomised hedonic (88%) and "tastes OK" (86%) measures. However, McNemar's test indicted significant bias in both cases (p<0.001), with a tendency for the VAS score of 0-50mm to underestimate the number of tastes-acceptable responses, compared to the hedonic and "tastes OK" assessments. As a result, alternative VAS cut-offs were explored, with 0-70mm selected to approximately match the cut-off on the hedonic scale where the neutral face becomes the negative. Using 70mm rather than 50mm in the VAS marginally increased the agreement with both the dichotomised hedonic and "tastes OK" measures to 91% but, more importantly, eliminated the previously observed bias (McNemar's test p=1.000, 0.683 respectively). All subsequent analysis used the cut-off of >70mm as a measure of unacceptable taste.

3.5. Associations with researcher observations and patient reported outcome measures

One patient did not have a record of facial expression/behaviours, hence they were excluded, and the analysis was based on n=620 cases. Associations between facial expressions and behaviours

(listed in Figure 2) were linked to cases where the medicines were reported to have an unacceptable taste, based on a composite outcome (Table 2). This was defined using the criteria defined previously. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available. In total, 255/620 (41%) of assessments identified the taste of medicines to be unacceptable. The associations between this outcome, and the various facial expressions and behaviours that were recorded were then assessed.

Table 2 shows the incidence of the facial expression or behaviour as a fraction of the total population. The behaviours are listed in order of Kendall's tau correlation coefficients, with those behaviours most strongly associated with unacceptable taste having the highest value of tau. Based on this metric, voicing disgust was the behaviour most strongly associated with negative taste. Of the 105 patients that voiced disgust, 95 (90%) reported the taste of the medicine to be negative. This rate was almost three-fold higher than the 160/515 (31%) who reported negative taste after not voicing disgust. Ordering the data in this way puts 'vomits' in last place, despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable. Since so few children vomited (n=7), the proportion of the total number of children who identified the taste as unacceptable and also displayed this behaviour (i.e. the sensitivity) was low, at 3%. Hence, although a vomiting child is almost certain to find the taste unacceptable, using this behaviour in isolation as a predictor of negative taste would miss the vast majority of patients who reported taste to be negative

Inter-rater agreement assessed via the use of short films and images were mixed; prevalent expressions were detected in >95% of cases, whereas some mild expressions were only detected in 40-50% of those viewing the images.

3.6. Analysis of medicine-specific taste assessment

Fifty-seven different drugs were observed in this study and the six most commonly administered were paracetamol, ibuprofen, prednisolone, co-amoxiclav, amoxicillin and clarithromycin, which made up 76% (n=477) of the total data set.

Comparisons were made across the six most commonly prescribed drugs. Table 3 ranks these drugs in order of taste, and reports the means scores on the hedonic scale and VAS, as well as the proportion of patients answering "no" to the, "Did you think the medicine tasted OK?" question.

The drugs can be divided into three groups based on this data: clarithromycin and prednisolone were found to be the worst tasting; amoxicillin and co-amoxiclav were mid-range, and ibuprofen and paracetamol were the best tasting medicines. The effect of brand was also investigated and the data is presented in Supplementary material 3.

In addition to reports of taste, the proportion of children who refused, vomited or spat out the medicines, was calculated and classified as unable to "use a medicinal product as intended". In total, this occurred in 33/620 (5%) administrations, which are reported for the six most commonly used drugs in Table 4.

Clarithromycin was the most commonly not taken as intended and was also the drug most frequently identified as having unacceptable taste, based on the previously defined composite outcome. However, there was insufficient data to suggest that the taste of the medicine was directly related to the ability to take the medicine as intended. Children may vomit due to their underlying illness rather than as a direct result of the taste of their medicine.

4. Discussion

Few studies have categorised acceptability of the taste of medicines. The results within this study agree with previous reports, where a neutral to positive hedonic response indicates acceptable taste [14-17]; Sjovall et al (1984) compared two brands of penicillin and reported that the acceptable taste mean hedonic score was within the neutral to positive range and an unacceptable taste was in the negative range [18]. Children were free to ask for support in completing the PRO measured and we did not collect data on how many received help in this aspect; it would be of value to consider how many, particularly in the youngest age group received support.

4.1. Interpretation of facial expressions and behaviours

Voicing disgust was the behaviour most strongly associated with unacceptable taste whilst, perhaps counterintuitively, vomiting was the behaviour least strongly correlated with unacceptable taste. Despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable, only a small number of patients vomited (n=7). As a result, whilst vomiting was a highly specific predictor of unacceptable taste, it had very low sensitivity. A similar pattern was also observed for the other facial expressions and behaviours, with sensitivity ranging from 10%-51%, and specificity from 84%-99%. This can be interpreted as indicating that, whilst it was uncommon to observe these facial expressions and behaviours in patients who found the taste of the medicine acceptable, displaying facial expressions and behaviours was not a strong indicator of unacceptability.

The behaviours used to inform the researcher observations were not always clearly defined; for example the use of physical restraint was not explicitly stated and further work is required to better understand what physical restraint may be considered acceptable.

The explicit definition of an acceptable medicine being "an overall ability of the patient and caregiver (defined as 'user') to use a medicinal product as intended (or authorised)" (Kozarewicz, 2014), includes the patient/caregiver's ability to access the medicine and comply with packaging requirements and for this study to demonstrate that the medicine was swallowed without incident. In total, 33/620 (5%) of medicines were spat out or vomited, and therefore unacceptable within this in patient population. This demonstrates that, although some of the behaviour and expressions observed may link more strongly to a negative taste, they do not automatically mean that the medicine was unacceptable.

In future studies, observations should ensure that the medicine was taken as intended; this may require a simple tool to ensure that the dose was completely swallowed without spitting out or vomiting. There is no need to include additional observations, as these were not strongly correlated to patient reported outcomes on the taste of medicines.

4.2. Medicine-specific taste assessment

Our results correlate well with other reports in the literature on the taste of medicines. Both paracetamol and ibuprofen have previously been reported to demonstrate acceptable taste (mid-range in a 5 point hedonic scale) in paediatric populations [19, 20]. Co-amoxiclav and amoxicillin were reported to have acceptable taste to 57-67% of children within this study, which is consistent with previous literature, which has reported co-amoxiclav to be preferred to other antibiotics in terms of taste, including amoxicillin [21-27]. Previous work suggested that 60% of children accepted amoxicillin without problems and 63% of children liked the taste of co-amoxiclav [27]. Prednisolone and clarithromycin were the least liked medicines, again consistent with published data, with results showing predominantly negative scores in hedonic scales or the equivalent of >70mm in a VAS [28-30, 23, 31, 32, 3]. This study identified differences in acceptance of certain brands of medicines; however, there was insufficient sample numbers to undertake a detailed analysis of the impact of brand on taste and acceptance. A future study that explored brand preference would be of value to patients and those purchasing commonly used medicines.

Using our criteria, the proportion of patients classifying taste as acceptable for the most liked of the frequently prescribed medicines (ibuprofen and paracetamol) was over 70%; the proportion classifying taste as acceptable for the least liked medicines (prednisolone and clarithromycin) was less than 30%. This study provides some standardised values for medicines used within a UK population, the thresholds reported here provide guidance for future medicines development where in addition to taste scores the overall risk-benefit of the medicine will need to be considered.

5. Conclusions

This study has generated data on the taste of medicines commonly used in paediatric populations aged 2-16 years. The results of this study suggest that patient reported outcome measures are a reliable and valid assessment of the taste of children's medicines, for children aged from 2-16 years.

These assessment tools offer a mechanism to evaluate the taste of other medicines (either novel products and formulations or medicines used orally in an off-label or unlicensed manner) to generate comparative data on the taste of medicines.

The data from this study coupled with previous literature on the taste of medicines provides evidence to suggest criteria to demonstrate acceptability of taste of medicines.

Our results suggest that criteria to demonstrate acceptability of taste are: a mean VAS score of <70mm; a mean hedonic score of ≤3 (neutral or positive face) and a non-negative response to the "Tastes OK?" question. Pragmatically, there is no need to use all methods. As the hedonic scale was understood across the widest age range, this should be the first choice method going forwards.

It would be prudent to ensure that any new product exceeds these scores to demonstrate that it is likely to have acceptable taste in practice.

Acknowledgements

The authors would like to thank the NIHR Clinical Research Network: West Midlands – Young Person's Steering Group (YPSG) for their input and advice in the development, conduct and dissemination of this study.

All patients, families and researchers at the participating sites are acknowledged for their participation in this study.

Figure Legends

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

Figure 3. Hedonic and VAS score distribution

Table headings

Table 1. Spearman's rho correlation values between patient-reported outcome measures.

	"Tastes OK" vs. Hedonic	"Tastes OK" vs. VAS	Hedonic vs. VAS
Overall	***0.80***	***0.78***	***0.84***
Age (Years)			
2-4	0.77	0.68	0.76
5-9	0.83	0.85	0.90
10-16	0.80	0.81	0.86

Table 2. Relationship between facial expression or a behaviour and a patient report of an unacceptable tasting medicine

	Cases Where Taste was Reported as unacceptable (n=255)				
Behaviour	***Not Displayed***	***Displayed***	**Tau**	**Sens.**	**Spec.**
Voices disgust	160/515 (31%)	95/105 (90%)	0.45	37%	97%
Eyes squeezed	131/460 (28%)	124/160 (78%)	0.44	49%	90%
Nose Wrinkle	125/433 (29%)	130/187 (70%)	0.38	51%	84%
Voices resistance	184/539 (34%)	71/81 (88%)	0.37	28%	97%
Refusal	192/551 (35%)	63/69 (91%)	0.36	25%	98%
Pursed Lips	168/505 (33%)	87/115 (76%)	0.33	34%	92%
Cries/Screams	201/559 (36%)	54/61 (89%)	0.32	21%	98%
Brow bulge	149/463 (32%)	106/157 (68%)	0.31	42%	86%
Physical restraint	219/579 (38%)	36/41 (88%)	0.25	14%	99%
Cries	216/573 (38%)	39/47 (83%)	0.24	15%	98%
Spits out	229/590 (39%)	26/30 (87%)	0.21	10%	99%
Vomits	248/613 (40%)	7/7 (100%)	0.13	3%	100%

Table 3. Patient reported taste scores by medicine

	Hedonic Score		VAS Score		Tastes OK?	Composite Outcome
Drug	***Mean***	***% unacceptable***	***Mean (mm)***	***% unacceptable***	**(% "No")**	***% unacceptable***
Clarithromycin (n=26)	3.5	62%	70	58%	65%	77%
Prednisolone (n=86)	3.6	58%	68	54%	61%	70%
Amoxicillin (n=30)	2.7	37%	37	23%	33%	43%

Co-Amoxiclav (n=50)	2.5	30%	39	24%	29%	33%
Paracetamol (n=193)	2.4	22%	37	21%	21%	29%
Ibuprofen (n=92)	2.1	14%	27	12%	12%	20%

Table 4. Percentage unable to take medicine as intended

Medicine	% spat out/vomited the medicine	Composite Outcome *% unacceptable*
Clarithromycin (n= 26)	23%	77%
Amoxicillin (n=30)	13%	43%
Prednisolone (n= 86)	9%	70%
Co-amoxiclav (n= 50)	6%	33%
Ibuprofen (n= 92)	2%	20%
Paracetamol (n= 193)	1%	29%

References

1. EMA. Guideline on pharmaceutical development of medicines for paediatric use EMA/CHMP/QWP/805880/2012. London. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/07/WC500147002.pdf. Accessed 18/06/2017.

2. Mistry P, Batchelor H. Evidence of acceptability of oral paediatric medicines: A review. J Pharm Pharmacol. 2016. doi:10.1111/jphp.12610.

3. Venables R, Batchelor H, Hodson J, Stirling H, Marriott J. Determination of formulation factors that affect oral medicines acceptability in a domiciliary paediatric population. International journal of pharmaceutics. 2015;480(1-2):55-62. doi:10.1016/j.ijpharm.2015.01.023.

4. Davies EH, Tuleu C. Medicines for Children: A Matter of Taste. Journal of Pediatrics. 2008;153(5):599-604. doi:10.1016/j.jpeds.2008.06.030.

5. Mistry P, Batchelor H. Methodology Used to Assess Acceptability of Oral Pediatric Medicines: A Systematic Literature Search and Narrative Review. Pediatric Drugs. 2017:1-11. doi:10.1007/s40272-017-0223-7.

6. Matza LS, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM et al. Pediatric Patient-Reported Outcome Instruments for Research to Support Medical Product Labeling: Report of the ISPOR PRO Good Research Practices for the Assessment of Children and Adolescents Task Force. Value Health. 2013;16(4):461-79. doi:http://dx.doi.org/10.1016/j.jval.2013.04.004.

7. FDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health and Quality of Life Outcomes. 2006;4(1):79. doi:10.1186/1477-7525-4-79.

8. Mistry P, Batchelor H, Tibbins C. Design of methodological tools to assess acceptability of paediatric medicines. Int J Pharm. 2016;511(2):1143-4. doi:http://dx.doi.org/10.1016/j.ijpharm.2016.06.098.

9. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A Systematic Review of Faces Scales for the Self-report of Pain Intensity in Children. Pediatrics. 2010;126(5):e1168-e98. doi:10.1542/peds.2010-1609.

10. Peryam DR, Pilgrim FJ. Hedonic scale method of measuring food preferences. Food Technology. 1957;11:9-14.

11. Katz ER, Kellerman J, Siegel SE. Behavioral distress in children with cancer undergoing medical procedures: developmental considerations. J Consult Clin Psychol. 1980;48(3):356-65.

12. Blount RL, Loiselle KA. Behavioural assessment of pediatric pain. Pain Research & Management : The Journal of the Canadian Pain Society. 2009;14(1):47-52.

13. Forestell CA, Mennella JA. More than Just a Pretty Face: The Relationship between Infant's Temperament, Food Acceptance, and Mothers' Perceptions of their Enjoyment of Food. Appetite. 2012;58(3):1136-42. doi:10.1016/j.appet.2012.03.005.

14. Motte J, Pedespan JM, Sevestre M, Chiron C. Acceptabilité et tolérance du valproate de sodium, granules à libération prolongée, en monothérapie chez l'enfant épileptique à partir de trois ans. Archives de Pédiatrie. 2005;12(10):1533-9. doi:http://dx.doi.org/10.1016/j.arcped.2005.07.009.

15. Mulla H, Buck H, Price L, Parry A, Bell G, Skinner R. 'Acceptability' of a new oral suspension formulation of mercaptopurine in children with acute lymphoblastic leukaemia.

Journal of oncology pharmacy practice : official publication of the International Society of Oncology Pharmacy Practitioners. 2015. doi:10.1177/1078155215577808.
16. Schwartz RH. Enhancing children's satisfaction with antibiotic therapy: A taste study of several antibiotic suspensions. Current Therapeutic Research. 2000;61(8):570-81. doi:10.1016/s0011-393x(00)80039-9.
17. Thompson A, Reader S, Field E, Shephard A. Open-label taste-testing study to evaluate the acceptability of both strawberry-flavored and orange-flavored amylmetacresol/2,4-dichlorobenzyl alcohol throat lozenges in healthy children. Drugs in R and D. 2013;13(2):101-7.
18. Sjövall J, Fogh A, Huitfeldt B, Karlsson G, Nylén O. Methods for evaluating the taste of paediatric formulations in children: A comparison between the facial hedonic method and the patients' own spontaneous verbal judgement. European Journal of Pediatrics. 1984;141(4):243-7.
19. Reader S, Shaw H, Hails S. taste-testing study in healthy volunteers (children) to investigate children's preference for ibuprofen or placebo suspension. 2006;7(2).
20. Smith CJ, Sammons HM, Fakis A, Conroy S. A prospective study to assess the palatability of analgesic medicines in children. Journal of Advanced Nursing. 2013;69(3):655-63.
21. Angelilli ML, Toscani M, Matsui DM, Rieder MJ. Palatability of oral antibiotics among children in an urban primary care center. Arch Pediatr Adolesc Med. 2000;154(3):267-70.
22. Gooch WM, Philips A, Rhoades R, Rosenberg R, Schaten R, Starobin S. Comparison of the efficacy, safety and acceptability of cefixime and amoxicillin/clavulanate in acute otitis media. The Pediatric Infectious Disease Journal. 1997;16(2):S21-S4.
23. Matsui D, Lim R, Tschen T, Rieder MJ. Assessment of the palatability of β-lactamase-resistant antibiotics in children. Archives of Pediatrics and Adolescent Medicine. 1997;151(6):599-602.
24. Powers JL, Gooch Iii WM, Oddo LP. Comparison of the palatability of the oral suspension of cefdinir vs. amoxicillin/clavulanate potassium, cefprozil and azithromycin in pediatric patients. Pediatric Infectious Disease Journal. 2000;19(12 SUPPL.):S174-S80.
25. Toscani M, Drehobl M, Freed J, Stool S. A multicenter, randomized, comparative assessment in healthy pediatric volunteers of the palatability of oral antibiotics effective in the therapy of otitis media. Current Therapeutic Research - Clinical and Experimental. 2000;61(5):278-85.
26. Ubaka CM, Udeogaranya OP, Ezeugwu N. Palatability of oral paediatric antibiotics commonly prescribed in a Nigerian specialist hospital. European Journal of Hospital Pharmacy: Science and Practice. 2013;20(2):122-4.
27. Dagan R, Shvartzman P, Liss Z. Variation in acceptance of common oral antibiotic suspensions. Pediatr Infect Dis J. 1994;13(8):686-90.
28. Hames H, Seabrook JA, Matsui D, Rieder MJ, Joubert GI. A PALATABILITY STUDY OF A FLAVORED DEXAMETHASONE PREPARATION VERSUS PREDNISOLONE LIQUID IN CHILDREN WITH ASTHMA EXACERBATION IN A PEDIATRIC EMERGENCY DEPARTMENT. Canadian Journal of Clinical Pharmacology. 2008;15(1):e95-e8.
29. Isa JM, Wong GK, Teraoka SS, Sera MJ, Tsushima MM, Yamamoto LG. Parental pediatric corticosteroid preferences. The American Journal of Emergency Medicine. 2001;19(1):29-31. doi:http://dx.doi.org/10.1053/ajem.2001.20025.
30. Kim MK, Yen K, Redman RL, Nelson TJ, Brandos J, Hennes HM. Vomiting of liquid corticosteroids in children with asthma. Pediatric emergency care. 2006;22(6):397-401. doi:10.1097/01.pec.0000221338.44798.6a.

31. Powers JL. Properties of azithromycin that enhance the potential for compliance in children with upper respiratory tract infections. The Pediatric Infectious Disease Journal. 1996;15(9):30-7.

32. Stafford L, Hope ME, Janney EP, Ailakis JG. Comparison of pediatric steroid mixtures. Australian Journal of Hospital Pharmacy. 1998;28(4):246-9.

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

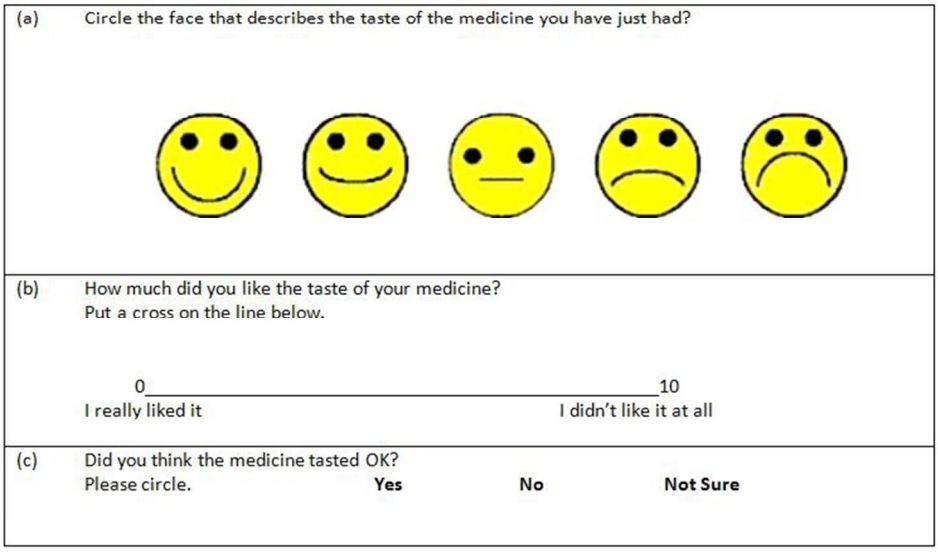


Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

67x53mm (300 x 300 DPI)

Facial expressions observed

Expression	Tick if observed prior to administration	Tick if observed during administration
Eyes squeezed shut or towards shut		
Brow bulge/lower (frown)		
Nose wrinkle		
Pursed lips		

Behaviours observed

Behaviour	Tick if observed prior to administration	Tick if observed during administration	
Child refuses medicine			=unacceptable
Child cries/screams			=unacceptable
Child requires physical restraint			=unacceptable
Child voices resistance			=unacceptable
	Tick if observed immediately after administration		
Child voices disgust			=unacceptable
Child vomits			=unacceptable
Child spits out medicine			=unacceptable
Child cries			=unacceptable

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

65x41mm (300 x 300 DPI)

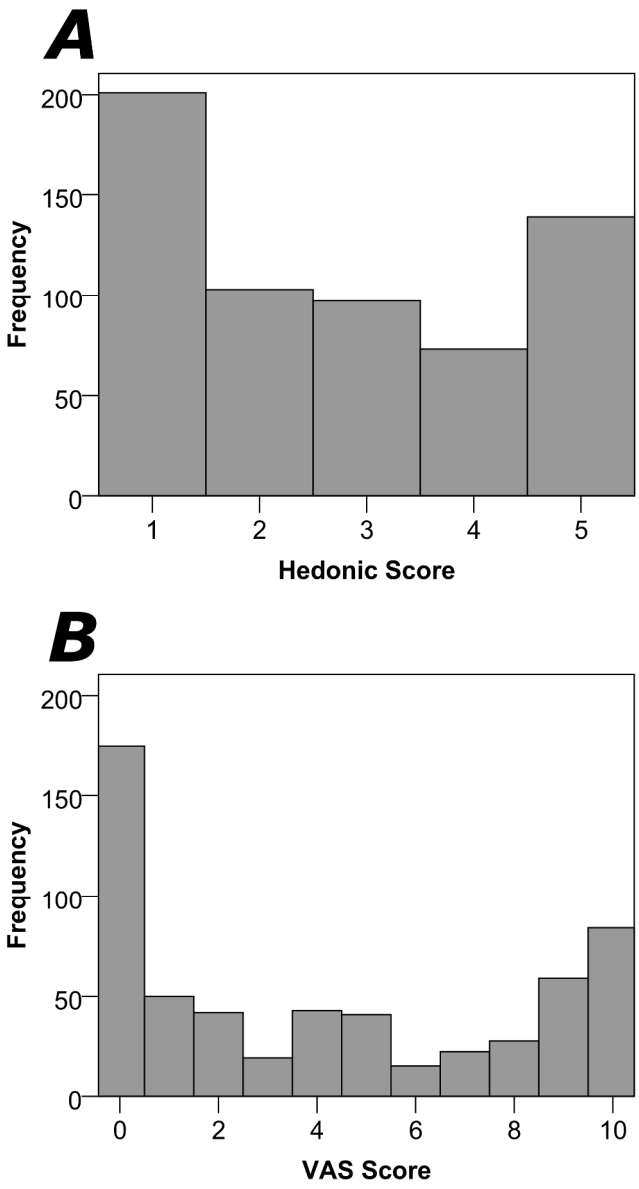
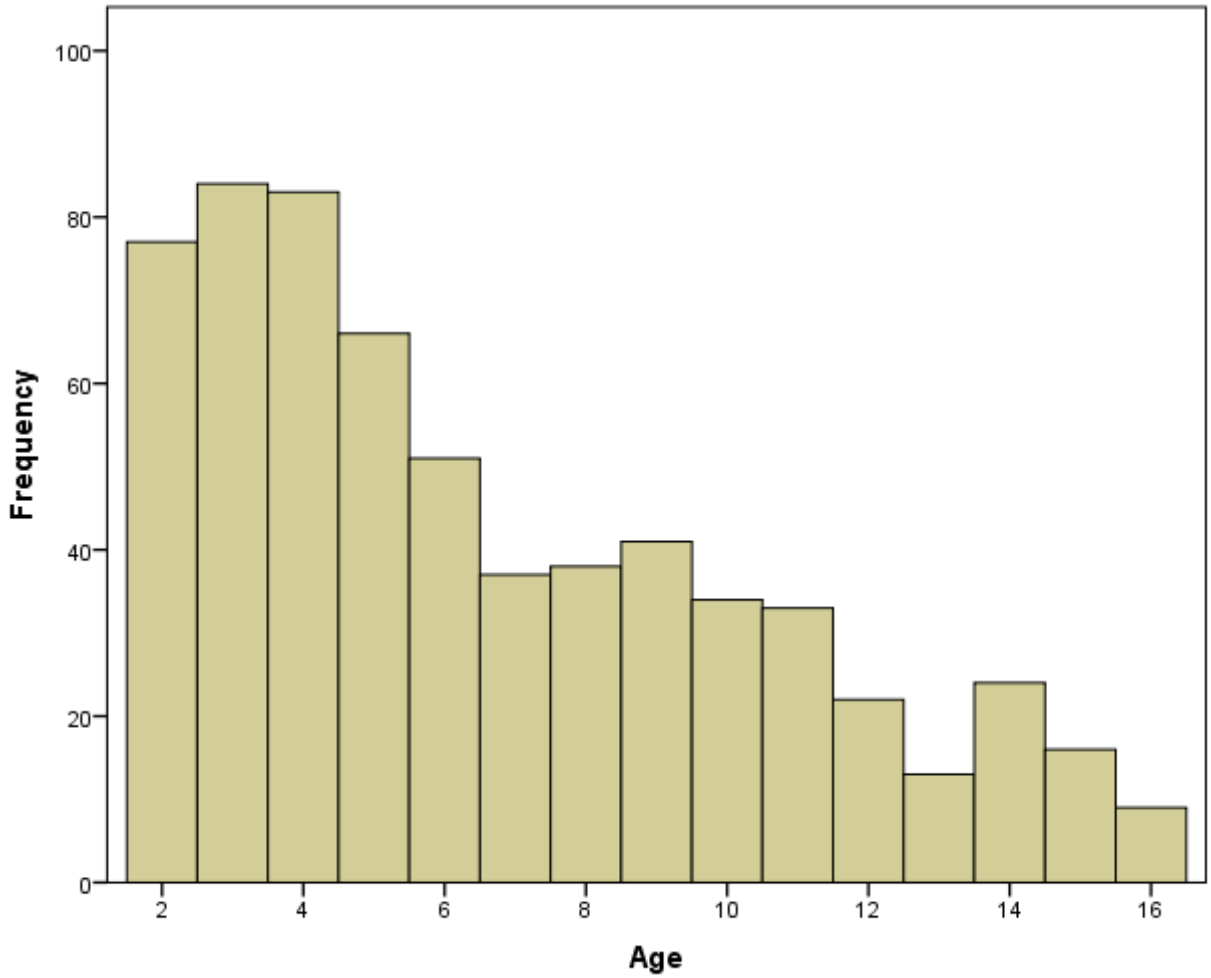
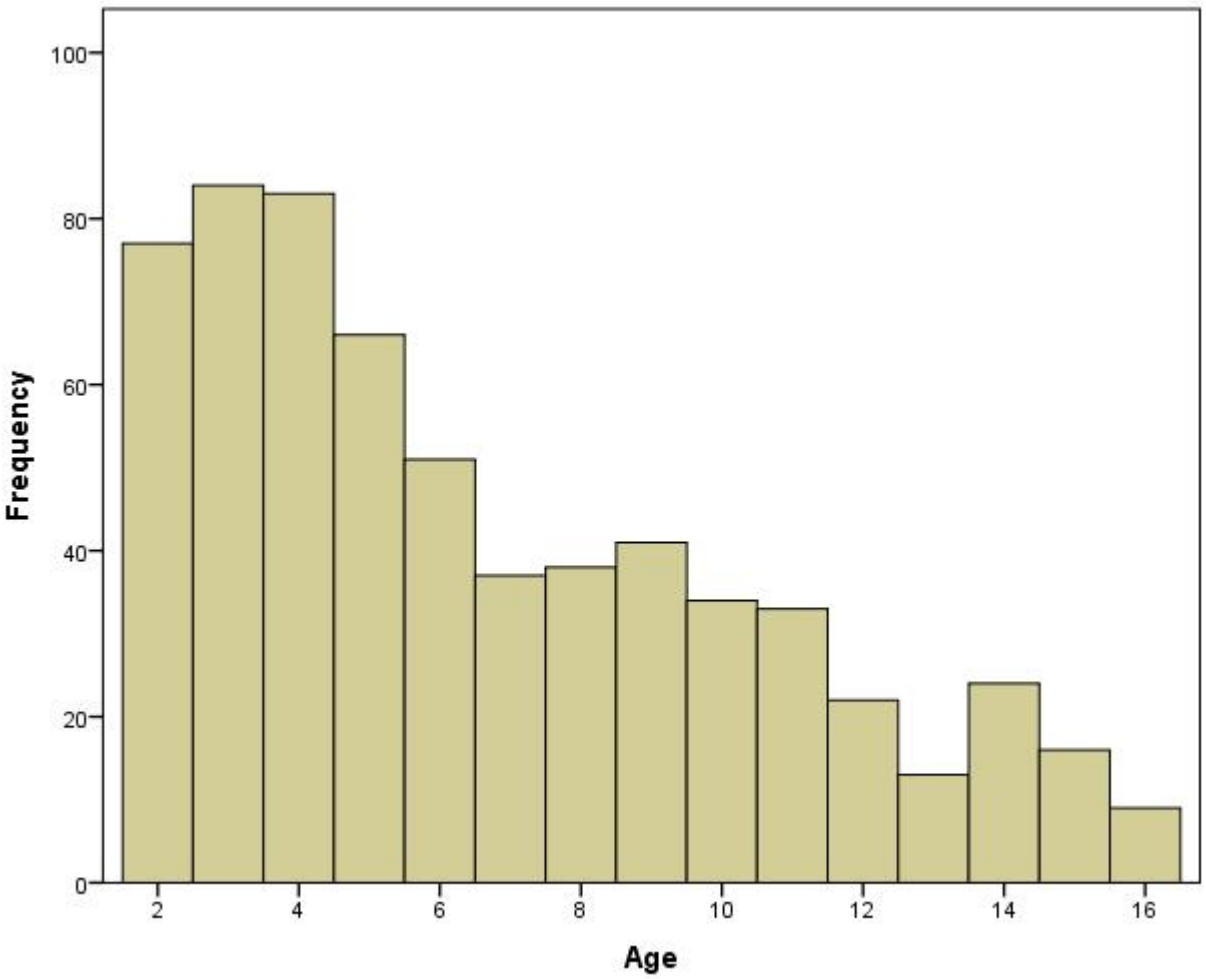


Figure 3. Hedonic and VAS score distribution

168x300mm (300 x 300 DPI)

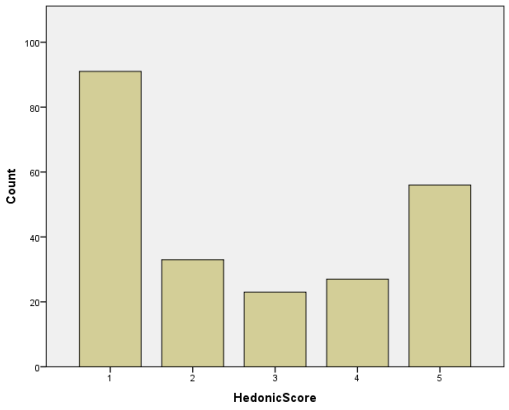
Supplementary Material 1. Distribution of participant age

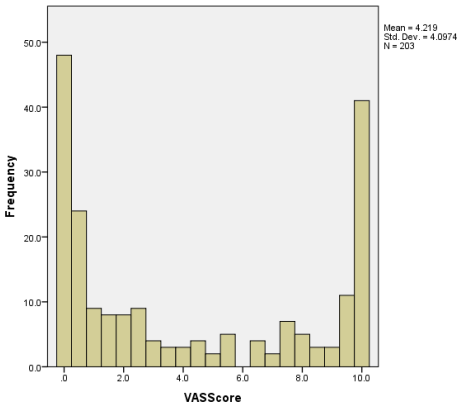




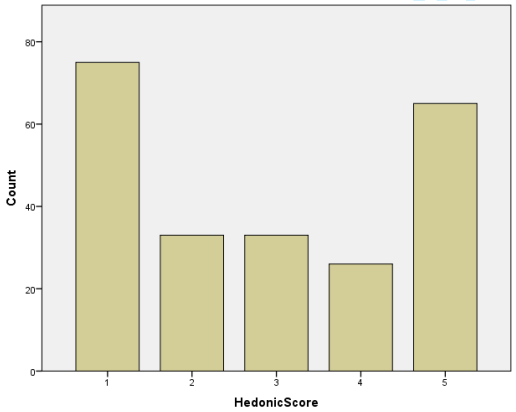
Supplementary Material 2. Age related distribution of responses from patient-reported assessment scales

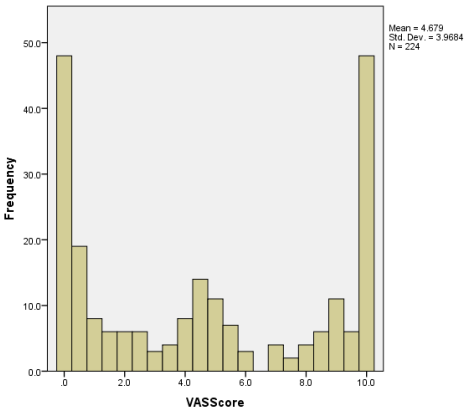
(a) Aged 2-4 years



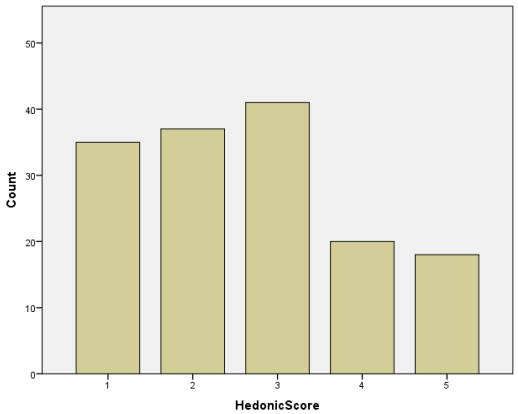


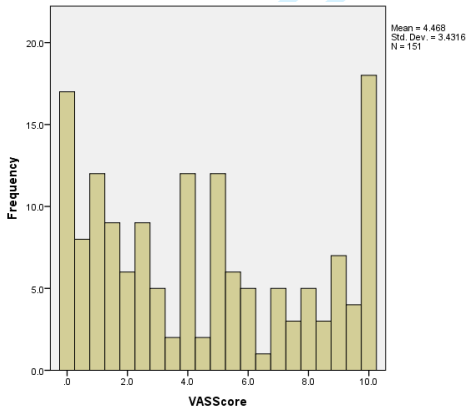
(b) Aged 5-9 years





(c) Aged 10-16 years





Supplementary material 3. Table of taste scores for the six most commonly administered drugs by brand

Drug/Manufacturer	N	Mean Hedonic Score	Mean VAS Score	Tastes OK (% "No")
Amoxicillin				
Athlone Laboratories	21	2.8	4.4	33%
Clamo	2	5.0	0.3	100%
Kent Pharmaceuticals	2	1.0	1.3	0%
Medreich PLC	1	3.0	10.0	100%
PL Holder	4	2.0	1.3	0%
Clarithromycin				
Abbott	11	3.7	7.4	82%
Klaricid	1	5.0	9.9	100%
MA Holder	1	2.0	4.2	0%
Sandoz Ltd.	13	3.4	6.7	54%
Co Amoxiclav				
Medreich PLC	1	3.0	5.0	0%
Rosemont Pharmaceutical Ltd	1	2.0	2.4	0%
Sandoz Ltd.	48	2.5	4.0	30%
Ibuprofen				
Fenpaed	9	2.1	3.4	22%
MA Holder	1	5.0	5.2	0%
McNeil Products	1	5.0	10.0	100%
Pinewood	78	2.0	2.5	10%
Reckitt Benckiser Healthcare	3	2.3	3.7	0%
Paracetamol				
Bristol Laboratories	1	3.0	4.0	100%
MA Holder	38	1.9	2.4	3%
McNeil Products	6	2.3	2.5	17%
Pinewood	113	2.6	4.3	28%
Rosemont Pharmaceutical Ltd	35	2.4	3.0	17%
Prednisolone				
Actavis	12	3.1	4.7	42%
Amdipharm	64	3.6	7.0	63%
Genetic SPA	1	3.0	5.2	0%
Teva UK Ltd	9	4.5	8.6	86%

STROBE Statement—checklist of items that should be included in reports of observational studies

This checklist was used and all items are provided within the manuscript: *Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines*

	Item No	Recommendation	Appears in manuscript (line number)
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	34-35
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Page 2 – Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	115-123
Objectives	3	State specific objectives, including any prespecified hypotheses	139-142
Methods			
Study design	4	Present key elements of study design early in the paper	145-150
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	167-180
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	168-171
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	183-208
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*172-174; 183-208*
Bias	9	Describe any efforts to address potential sources of bias	211-212
Study size	10	Explain how the study size was arrived at	216-218
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings	193-196; 219-232

		were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	220-222; 215-232
		(*b*) Describe any methods used to examine subgroups and interactions	240-241
		(*c*) Explain how missing data were addressed	227-228
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	222-224
		(*e*) Describe any sensitivity analyses	284-290

Continued on next page

Results			**Appears in manuscript**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	250-259
		(b) Give reasons for non-participation at each stage	250-259
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	237-239
		(b) Indicate number of participants with missing data for each variable of interest	250-259
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*250-259*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	261-364
		(*b*) Report category boundaries when continuous variables were categorized	301-313
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	343-364
Discussion			
Key results	18	Summarise key results with reference to study objectives	366-422
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	56-69
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	425-438
Generalisability	21	Discuss the generalisability (external validity) of the study results	433-438
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	90-95

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-021961.R1
Article Type:	Research
Date Submitted by the Author:	06-Apr-2018
Complete List of Authors:	Mistry, Punam; University of Birmingham, School of Pharmacy Stirling, Heather; University Hospitals Coventry and Warwickshire NHS Trust, Paediatrics Callens, Claire; NIHR Clinical Research Network:West Midlands – Young Person's Steering Group (YPSG), CRN: Children Hodson, James; Queen Elizabeth Hospital Birmingham, Institute of Translational Medicine Batchelor, Hannah; University of Birmingham, School of Pharmacy
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Patient-centred medicine, Pharmacology and therapeutics
Keywords:	THERAPEUTICS, PAEDIATRICS, palatability, ORAL MEDICINE

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines

Punam Mistry[1], Heather Stirling[2], Claire Callens[3], James Hodson[4] and Hannah Batchelor[1] on behalf of SPaeDD-UK project

(Smart Paediatric Drug Development – UK, accelerating paediatric formulation development http://www.paediatricscienceuk.com)

[1] School of Pharmacy, Institute of Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, UK

[2] University Hospital Coventry and Warwickshire NHS Trust, Clifford Bridge Road, Coventry, CV2 2DX, UK

[3] NIHR Clinical Research Network West Midlands, Institute of Research and Development, Birmingham Research Park, Vincent Drive, Birmingham, B15 2SQ, UK

[4] Institute of Translational Medicine, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston, Birmingham, B15 2GW, UK

Affiliation: Pharmacy and Therapeutics, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT, United Kingdom

Corresponding author: Hannah Batchelor, School of Pharmacy, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT
Email: h.k.batchelor@bham.ac.uk
Phone: +44 (0)121 414 3717

Abstract

Objective: To evaluate the age-appropriateness and suitability of patient reported outcome measures to assess the acceptability of the taste of oral liquid medicines in children.

Design and setting: An observational mixed methods study involving children aged 2-16 years old taking oral liquid medicine in paediatric inpatient wards across the West Midlands (UK). Assessment tools included: patient-reported scores on the taste of medicines via a 5-point facial hedonic scale; a visual analogue scale (VAS); a question, "Did you think the medicine tasted OK?" and researcher observations of facial expressions and behaviours immediately before, during and after administration.

Results: 611 children participated. The percent unable to complete the scales were 7% (n=46) for the VAS; 2% (n=15) for the hedonic scale and 1% (n=7) for the question about taste. Significant correlations (Spearman's rho) were observed between the patient reported outcome measures: 0.80 and 0.78 for the taste question and hedonic and VAS respectively and 0.84 for the hedonic and VAS. Researcher observations demonstrated the ability of the patient to take the medicine as intended but did not provide sensitive measures of taste. 5% of administrations were not taken as intended by the children. Medicines known to have poor taste (clarithromycin and prednisolone) showed mean hedonic and VAS scores of ≥3.5 and >65mm respectively.

Conclusions: Patient reported outcome measures correlate with each other and are a useful means to assess the taste (and acceptability) of medicines. Hedonic scales are better understood by children and should be the first choice tool in the assessment of medicines taste.

Key words: medicines, taste, acceptability, age-appropriate, hedonic scale, paediatric medicine, patient-reported outcome measures, VAS

Strengths and limitations of this study

- This is the first study to compare methodologies to assess the acceptability of taste of liquid medicines in a large UK based paediatric population aged 2-16 years
- The sample size in this study provided large data sets of six key medicines, which provides a representative comparison for newly developed products
- This study was conducted within an inpatient environment and the acceptance of taste of medicines in a domiciliary environment may differ
- The study design captured the most relevant aspects of acceptability of taste whilst minimising the burden to participants, it was not possible to measure every aspect. Suggestions for future research include measurement of: impact of the devices used to administer medicines; use of alternative hedonic scales (e.g. those with 2-9 faces); alternative anchor phrases; an endpoint of neutral and not positive within a hedonic scale; further exploration of medicines that tasted OK as well as those with a reported negative taste.

What is already known about this subject?

- New medicines for children must be demonstrated to be acceptable to a paediatric population
- Measurement of acceptability of the taste of medicines is complex and a wide range of methods have been used previously, making comparison between studies complex.
- There is a need for age-appropriate reproducible and reliable methods to measure the acceptability of medicines

What this study adds

- Patient reported outcome measures offer a pragmatic means to assess the taste of medicines in children aged from 2-16 years

- 5-point hedonic scales were better understood compared to visual analogue scales in children aged 2-16 years
- Although 41% of medicines were reported to have unacceptable taste only 5% were so bad that they could not be taken as intended

Funding Statement

This work was conducted as part of the SPaeDD-UK project (Smart Paediatric Drug Development – UK, a project co-funded by Innovate UK and the contributing companies of AstraZeneca, Bristol Myers Squibb, GlaxoSmithKline, Juniper Pharmaceuticals and Pfizer. (http://www.paediatricscienceuk.com).

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Data sharing statement

Additional data is available in the Supplementary files. The full data set is held by the corresponding author, please email with any requests for extra data.

Authorship contributions

Punam Mistry contributed to the acquisition, analysis and interpretation of the data.

Heather Stirling contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Claire Callens contributed to the design of the study and acquisition of the data

James Hodson contributed to the design, statistical analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Hannah Batchelor contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments. She is the corresponding author for this work

Word count of manuscript body: 3836

1. Introduction

Measurement of acceptability of paediatric medicines is important for two major reasons: (i) poor acceptability is likely to be associated with poor compliance and therefore suboptimal therapy; (ii) regulatory guidelines dictate that paediatric medicines' acceptability should be assessed (preferably in children) [1]. Patient acceptability has been defined as, "an overall ability of the patient and caregiver (defined as 'user') to use a medicinal product as intended (or authorised)" [1]. A major barrier in the assessment of acceptability and subsequent interpretation of results is the lack of guidance on how an acceptability test should be conducted and the criteria used to determine acceptability in children [2]. It is unlikely that a single acceptability test will be valid for all medicines, due to the need to consider the benefit:risk balance on an individual basis.

Taste was previously identified as the biggest barrier in adherence and compliance in a paediatric population [3]. Previous techniques to evaluate acceptability focussed on taste and were based on methodology adapted from sensory analysis undertaken in the food industry. A variety of scales have been used to measure the acceptability of medicines, yet it remains unclear whether any of the scales is better for a particular purpose with regard to validity, reliability, feasibility, and preference [4, 5]. Patient reported measures of taste offer a pragmatic means to collect data on products. However, there is often concern about the reliability of these measures, particularly in a paediatric population [6]. Sensory analysis is commonly used to show a preference for one product over another, but this is not always relevant to medicines as often there is no alternative available. In addition, it can be rationalized that the therapeutic benefits of medicines outweigh the negative taste. It is complex to distinguish between acceptability and likeability of a medicine's taste, yet this difference may be critical for appropriate interpretation of resulting data.

The validity of any new method relies on how well it compares to other measures [7]. The development of a reliable method to assess the taste of medicines requires comparison of a variety of tools. This study compares a range of methods to assess the taste of medicines within a large

paediatric inpatient population. Independent researcher observations are also compared to the self-reported data. The results of this study will be used to propose a suitable method that can be used for future taste assessments.

2. Materials and Methods

Three patient-reported outcome (PRO) measures were compared to each other, and to researcher observations of medicines administration in an observational mixed methods study.

2.1. Patient and Public Involvement

Bespoke PRO tools were developed for this study based on previous methodologies and in consultation with the National Institute for Health Research (NIHR) Children Specialty's Young Person's Advisory Group (West Midlands) [8]. The young people (aged 11-18 years) reviewed the tools and provided feedback that the tools were age-appropriate. The same young people provided feedback on the trial materials including information sheets and how to minimise the burden to participants during the conduct of the study. The results are available to participants as a poster summary from the corresponding author's personal webpage (www.hannahbatchelor.com); this poster was also reviewed by the young person's group.

2.2. Patient reported outcome measures used

The hedonic scale selected was a genderless image where the mouth was the only expressive facial feature (see Figure 1(a)). The images were yellow to ensure they stood out from the background paper. They were obtained from S-cool the revision website (http://www.s-cool.co.uk/gcse/food-technology/systems-and-control/revise-it/sensory-evaluation (accessed December 2015)). Children and young people preferred simple faces and felt that this would be most appropriate for the youngest age group [8]; anchor phrases were not included on the hedonic scale to keep this tool basic. The visual analogue scale (VAS) included anchor phrases selected by the young persons'

group as the most clear and relevant [8], these were used at the extreme ends of the continuous scale.

The direction of change was from positive to negative, which corresponds to the extensive data on hedonic scales used for pain assessment in children [9]. Previous studies in adult studies of foods showed no difference based on structural variations that read from positive to negative or vice versa [10].

The third PRO was a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure.

2.3. Participants and Setting

Paediatric inpatients aged between 2 to 16 years taking any oral liquid medicine as part of their medical care were recruited using convenience sampling, from inpatient wards at 11 sites across the West Midlands. The study was conducted between December 2015- April 2016. Ethical approval was granted by London Surrey borders ethical committee (REC reference: 15/LO/1253; IRAS project ID: 179684).

Demographic information was obtained on participant's age and whether this was their first dose of the medicine under evaluation. The medicine name; brand; manufacturer; concentration (strength); dose administered and product batch number was recorded.

Each participant was observed by a researcher as they took their medicine. Some medicines were provided to the patient as an oral liquid following extemporaneous preparation within the clinical setting, for example, dispersion of prednisolone tablets immediately prior to dosing. Depending on the capability of the child, the medicines were either self-administered, or administered by nursing staff and/or parents. Participants were asked not to mix the medicine with any other food product, as this might influence the participant's responses.

2.4. Patient-reported outcome tools

Participants were asked to complete (a) 5-point facial hedonic scale and (b) 10cm VAS (Figure 1) immediately after administration of their medicine; both scales were provided on separate paper documents in a randomized order. Children were free to ask for support in completing the questionnaires from parents, nursing staff or the researcher present. The cognitive function of children was not assessed and age may not always predict a child's ability to complete the questionnaire, therefore all children were free to ask for support if requried. Both reporting documents included a third PRO (Figure 1 (c)) as a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure. The purpose behind this question was to endorse the reliability of the participant's reporting from both scale-based questionnaires.

Figure 1

The data from the scales (a and b) were transformed into numbers for subsequent statistical analysis. Hedonic expressions were converted to scores of 1 to 5 with 1 being the most positive expression and 5 the most negative. VAS scores were reported as measurements from the extreme left hand side of the scale in mm; this provided scores of 0-100 mm (although scale was written as 0-10 cm).

2.5. Researcher Observations

A facial expression and behavioural scale to capture researcher observations was a 12-point tick chart designed in-house (Figure 2), based on an existing scale used to measure behavioural distress in children undergoing medical procedures [11], the Procedural Behavior Rating Scale (PBRS). The PBRS was developed specifically for infants and children with cancer, who were undergoing bone marrow aspirations or lumbar punctures. It contained 13 behavioural codes including: cry, cling, pain verbal, flail, stall, scream, refuse position, restrain, emotional support, muscular rigidity and requests termination [11]. The in-house tool was based on this scale plus other scales used to measure behavioural distress in children [12]. The facial expressions included on the scale were derived from

previous studies that assessed food-liking in children based on their facial expressions; typically negative tastes are associated with eyes squeezing, brow bulge, nose wrinkle and pursed lips [13].

Figure 2

A series of nine short films or still pictures of children were made available to researchers participating in the study to assess the inter-rater agreement in the facial expressions displayed.

2.6. Statistical analysis

A sample size was not fixed for this study at the outset as there was no appropriate power calculation. Advice from a statistician was to ensure we had sufficient numbers (<50) paired data to ensure that selectivity and sensitivity of the methods could be demonstrated.

Initially, the mean age of participants that were unable to understand the assessments were compared to those who could using Mann-Whitney tests. The same approach was also used to compare the scores for patients receiving their first dose, relative to those who had previously received the medicine. Age was then divided into categories, and the proportions of participants scoring in the extreme categories for the scales were compared using Fisher's exact tests.

Spearman's correlation coefficients were used to assess the degree of correlation between the assessments for the cohort as a whole, and within each of the age categories. In this analysis, the "Did you think the medicine tasted OK?" question was treated as an ordinal scale, with the response of Don't Know being between Yes and No. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available

The assessments were then dichotomised, and compared using McNemar's tests to assess for marginal homogeneity in the 2x2 tables, and Kappa to assess agreement. The three assessments were then combined into a composite score, which was compared with reported behaviours using

Kendall's tau correlation coefficients. All analyses were performed using IBM SPSS 22 (IBM Corp. Armonk, NY), and $p<0.05$ deemed to be indicative of statistical significance throughout.

3. Results and discussion

Data were available for 628 administrations to 611 children aged between 2-16 years. The median participant age was 6 years. Further details on the distribution of the participant ages can be found in supplementary material 1.

To ease analysis of data, the population for this study was stratified by age into three groups: 2-4 years (n=237); 5-9 years (n=227) and 10-16 years (n=147).

The medicine was administered as the first dose in 162 cases. There was no evidence of a significant difference in the hedonic or VAS scores between those receiving their first dose of a medicine, compared to those who had previous administrations (p=0.336, 0.909 respectively). For all subsequent analysis the data was pooled for those receiving their first and subsequent doses of medicine.

3.1. Completeness of patient-reported assessment scales

The assessment scales were not completed by all of the study participants. The VAS had the lowest completion rate, where 46 (7%) were not completed due to lack of understanding by the child, compared to 15 (2%) for the hedonic scale. There were 7 (1%) cases where the child did not understand the question, "Did you think the medicine taste OK?". The range and mean age of those unable to complete the scales were 2-5 years with a mean of 2.4 years for the hedonic; 2-6 years with a mean of 3.1 years for the VAS and 2-3 years with a mean of 2.3 years for the taste question. In each case, participants unable understand the assessment methods were significantly younger than the remainder of the cohort ($p<0.001$ each assessment). The cognitive function of children was not

assessed within this study and there was an assumption of cognitive normal for age for all participants.

3.2. Distribution of responses to patient-reported assessment scales

In the patient reported scales (VAS and hedonic), the responses tended to be biased towards the extreme responses, with 56% of responses being in the highest or lowest categories for hedonic score (Figure 3a) and, 55% within 1cm from either end of the scale in the VAS (Figure 3b).

Figure 3

The use of the extreme ends of the scales was greater in the younger populations ($p<0.001$) (data is shown in Supplementary material 2). For those that completed the hedonic scale, 64% of those aged 2-4 years used the extreme ends of the scale compared to 60% of those aged 5-9; 35% of those aged 10-16 years. A similar trend was observed in the VAS data where 66% of those aged 2-4 years were within 10mm of either end of the scale compared to 60% of 5-9 year olds and 39% of 10-16 year olds.

3.3. Reliability of assessment, "Did you think the medicine tasted OK?"

The question, "Did you think the medicine tasted OK?" was asked to participants on two occasions for each medicine administration; once following the hedonic and once following the VAS (which were presented in a randomized order). In the 600 cases where participants responded to both questions, the same response was given in 569 (95%) of cases, giving a Kappa statistic of 0.91. In cases where the answers were not consistent, the median age of the respondents was 5 years (mean age = 6.3 years). For subsequent analysis, the two versions of this question were merged, with discrepancies resolved by taking the most positive response given by the participant.

3.4. Correlation between patient reported outcome measures

Significant correlations were observed between the hedonic scale score, VAS and "Did you think the medicines tasted OK?" question (all p<0.001 Table 1), with the strongest correlation observed between the hedonic and VAS scores (Spearman's rho=0.84). The weakest correlations were consistently observed in the youngest patients (age 2-4 years), implying that this group of patients had the lowest consistency in scores given across the different assessments. However, despite this, the consistency between the scores was still reasonable, with correlation coefficients ranging from 0.68- 0.77.

The assessments were dichotomised, to identify those responses that classified a medicine as tasting acceptable. For the hedonic scale, this was defined as a neutral to positive response (the leftmost three faces of Figure 1a). For the *"Did you think the medicine tasted OK?"* question, grouping the "yes" and "not sure" responses was considered most appropriate, as the overall purpose is the assessment of acceptability of taste of a medicine; the response "no" is clearly aligned to an unacceptable taste, making other responses acceptable. The level of agreement between the resulting dichotomised scores was high (90%), with a Kappa statistic of 0.782, and no indication of bias (McNemar's test: p=0.519).

Correlation of the hedonic scale to the "tastes OK" question is simple; however interpretation of the VAS is more complex. The VAS had anchor phrases at either end to maintain the continuous scale, so it was not obvious where neutral or positive was ranked on the scale. Initially, an arbitrary cut-off of <50mm in the VAS was used to define acceptability, which gave reasonable agreement with both the dichotomised hedonic (88%) and "tastes OK" (86%) measures. However, McNemar's test indicted significant bias in both cases (p<0.001), with a tendency for the VAS score of 0-50mm to underestimate the number of tastes-acceptable responses, compared to the hedonic and "tastes OK" assessments. As a result, alternative VAS cut-offs were explored, with 0-70mm selected to approximately match the cut-off on the hedonic scale where the neutral face becomes the negative.

Using 70mm rather than 50mm in the VAS marginally increased the agreement with both the dichotomised hedonic and "tastes OK" measures to 91% but, more importantly, eliminated the previously observed bias (McNemar's test p=1.000, 0.683 respectively). All subsequent analysis used the cut-off of >70mm as a measure of unacceptable taste.

3.5. Associations with researcher observations and patient reported outcome measures

One patient did not have a record of facial expression/behaviours, hence they were excluded, and the analysis was based on n=620 cases. Associations between facial expressions and behaviours (listed in Figure 2) were linked to cases where the medicines were reported to have an unacceptable taste, based on a composite outcome (Table 2). This was defined using the criteria defined previously. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available. In total, 255/620 (41%) of assessments identified the taste of medicines to be unacceptable. The associations between this outcome, and the various facial expressions and behaviours that were recorded were then assessed.

Table 2 shows the incidence of the facial expression or behaviour as a fraction of the total population. The behaviours are listed in order of Kendall's tau correlation coefficients, with those behaviours most strongly associated with unacceptable taste having the highest value of tau. Based on this metric, voicing disgust was the behaviour most strongly associated with negative taste. Of the 105 patients that voiced disgust, 95 (90%) reported the taste of the medicine to be negative. This rate was almost three-fold higher than the 160/515 (31%) who reported negative taste after not voicing disgust. Ordering the data in this way puts 'vomits' in last place, despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable. Since so few children vomited (n=7), the proportion of the total number of children who identified the taste as unacceptable and also displayed this behaviour (i.e. the sensitivity) was low, at 3%. Hence, although a vomiting child is almost certain to find the taste unacceptable, using this behaviour in isolation as a predictor of negative taste would miss the vast majority of patients who reported taste to be negative

Inter-rater agreement assessed via the use of short films and images were mixed; prevalent expressions were detected in >95% of cases, whereas some mild expressions were only detected in 40-50% of those viewing the images.

3.6. Analysis of medicine-specific taste assessment

Fifty-seven different drugs were observed in this study and the six most commonly administered were paracetamol, ibuprofen, prednisolone, co-amoxiclav, amoxicillin and clarithromycin, which made up 76% (n=477) of the total data set.

Comparisons were made across the six most commonly prescribed drugs. Table 3 ranks these drugs in order of taste, and reports the means scores on the hedonic scale and VAS, as well as the proportion of patients answering "no" to the, "Did you think the medicine tasted OK?" question.

The drugs can be divided into three groups based on this data: clarithromycin and prednisolone were found to be the worst tasting; amoxicillin and co-amoxiclav were mid-range, and ibuprofen and paracetamol were the best tasting medicines. The effect of brand was also investigated and the data is presented in Supplementary material 3.

In addition to reports of taste, the proportion of children who refused, vomited or spat out the medicines, was calculated and classified as unable to "use a medicinal product as intended". In total, this occurred in 33/620 (5%) administrations, which are reported for the six most commonly used drugs in Table 4.

Clarithromycin was the most commonly not taken as intended and was also the drug most frequently identified as having unacceptable taste, based on the previously defined composite outcome. However, there was insufficient data to suggest that the taste of the medicine was directly related to the ability to take the medicine as intended. Children may vomit due to their underlying illness rather than as a direct result of the taste of their medicine.

4. Discussion

Few studies have categorised acceptability of the taste of medicines. The results within this study agree with previous reports, where a neutral to positive hedonic response indicates acceptable taste [14-17]; Sjovall et al (1984) compared two brands of penicillin and reported that the acceptable taste mean hedonic score was within the neutral to positive range and an unacceptable taste was in the negative range [18]. Children were free to ask for support in completing the PRO measured and we did not collect data on how many received help in this aspect; it would be of value to consider how many, particularly in the youngest age group received support. Many of the children aged 2-5 years were able to provide reliable data on the taste of medicines demonstrating that the scales and questions used within this study are suitable for very young participants.

4.1. Interpretation of facial expressions and behaviours

Voicing disgust was the behaviour most strongly associated with unacceptable taste whilst, perhaps counterintuitively, vomiting was the behaviour least strongly correlated with unacceptable taste. Despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable, only a small number of patients vomited (n=7). As a result, whilst vomiting was a highly specific predictor of unacceptable taste, it had very low sensitivity. A similar pattern was also observed for the other facial expressions and behaviours, with sensitivity ranging from 10%-51%, and specificity from 84%-99%. This can be interpreted as indicating that, whilst it was uncommon to observe these facial expressions and behaviours in patients who found the taste of the medicine acceptable, displaying facial expressions and behaviours was not a strong indicator of unacceptability.

The behaviours used to inform the researcher observations were not always clearly defined; for example the use of physical restraint was not explicitly stated and further work is required to better understand what physical restraint may be considered acceptable.

The explicit definition of an acceptable medicine being “an overall ability of the patient and caregiver (defined as ‘user’) to use a medicinal product as intended (or authorised)” (Kozarewicz, 2014), includes the patient/caregiver’s ability to access the medicine and comply with packaging requirements and for this study to demonstrate that the medicine was swallowed without incident. In total, 33/620 (5%) of medicines were spat out or vomited, and therefore unacceptable within this in patient population. This demonstrates that, although some of the behaviour and expressions observed may link more strongly to a negative taste, they do not automatically mean that the medicine was unacceptable.

In future studies, observations should ensure that the medicine was taken as intended; this may require a simple tool to ensure that the dose was completely swallowed without spitting out or

vomiting. There is no need to include additional observations, as these were not strongly correlated to patient reported outcomes on the taste of medicines.

4.2. Medicine-specific taste assessment

Our results correlate well with other reports in the literature on the taste of medicines. Both paracetamol and ibuprofen have previously been reported to demonstrate acceptable taste (mid-range in a 5 point hedonic scale) in paediatric populations [19, 20]. Co-amoxiclav and amoxicillin were reported to have acceptable taste to 57-67% of children within this study, which is consistent with previous literature, which has reported co-amoxiclav to be preferred to other antibiotics in terms of taste, including amoxicillin [21-27]. Previous work suggested that 60% of children accepted amoxicillin without problems and 63% of children liked the taste of co-amoxiclav [27]. Prednisolone and clarithromycin were the least liked medicines, again consistent with published data, with results showing predominantly negative scores in hedonic scales or the equivalent of >70mm in a VAS [28-30, 23, 31, 32, 3]. This study identified differences in acceptance of certain brands of medicines; however, there was insufficient sample numbers to undertake a detailed analysis of the impact of brand on taste and acceptance. A future study that explored brand preference would be of value to patients and those purchasing commonly used medicines.

Using our criteria, the proportion of patients classifying taste as acceptable for the most liked of the frequently prescribed medicines (ibuprofen and paracetamol) was over 70%; the proportion classifying taste as acceptable for the least liked medicines (prednisolone and clarithromycin) was less than 30%. This study provides some standardised values for medicines used within a UK population, the thresholds reported here provide guidance for future medicines development where in addition to taste scores the overall risk-benefit of the medicine will need to be considered.

4.3. Recommended tools to assess acceptability

This study has correlated three simple patient reported measures of medicines taste acceptability. It has also provided comparative data from existing medicines. Regulations mandate that all new medicines need to be demonstrated to be acceptable to children [1]. This study provides pragmatic and reliable tools to conduct this assessment. Furthermore, comparison of the results from a new medicine using these tools can be directly compared to existing medicines to support evidence of acceptance.

5. Conclusions

This study has generated data on the taste of medicines commonly used in paediatric populations aged 2-16 years. The results of this study suggest that patient reported outcome measures are a reliable and valid assessment of the taste of children's medicines, for children aged from 2-16 years. These assessment tools offer a mechanism to evaluate the taste of other medicines (either novel products and formulations or medicines used orally in an off-label or unlicensed manner) to generate comparative data on the taste of medicines.

The data from this study coupled with previous literature on the taste of medicines provides evidence to suggest criteria to demonstrate acceptability of taste of medicines.

Our results suggest that criteria to demonstrate acceptability of taste are: a mean VAS score of <70mm; a mean hedonic score of ≤3 (neutral or positive face) and a non-negative response to the "Tastes OK?" question. Pragmatically, there is no need to use all methods. As the hedonic scale was understood across the widest age range, this should be the first choice method going forwards.

It would be prudent to ensure that any new product exceeds these scores to demonstrate that it is likely to have acceptable taste in practice.

Acknowledgements

The authors would like to thank the NIHR Clinical Research Network: West Midlands – Young Person's Steering Group (YPSG) for their input and advice in the development, conduct and dissemination of this study.

All patients, families and researchers at the participating sites are acknowledged for their participation in this study.

Figure Legends

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

Figure 3. Hedonic and VAS score distribution

Table headings

Table 1. Spearman's rho correlation values between patient-reported outcome measures.

	"Tastes OK" vs. Hedonic	"Tastes OK" vs. VAS	Hedonic vs. VAS
Overall	***0.80***	***0.78***	***0.84***
Age (Years)			
2-4	0.77	0.68	0.76
5-9	0.83	0.85	0.90
10-16	0.80	0.81	0.86

Table 2. Relationship between facial expression or a behaviour and a patient report of an unacceptable tasting medicine

	Cases Where Taste was Reported as unacceptable (n=255)				
Behaviour	***Not Displayed***	***Displayed***	**Tau**	**Sens.**	**Spec.**
Voices disgust	160/515 (31%)	95/105 (90%)	0.45	37%	97%

Eyes squeezed	131/460 (28%)	124/160 (78%)	0.44	49%	90%
Nose Wrinkle	125/433 (29%)	130/187 (70%)	0.38	51%	84%
Voices resistance	184/539 (34%)	71/81 (88%)	0.37	28%	97%
Refusal	192/551 (35%)	63/69 (91%)	0.36	25%	98%
Pursed Lips	168/505 (33%)	87/115 (76%)	0.33	34%	92%
Cries/Screams	201/559 (36%)	54/61 (89%)	0.32	21%	98%
Brow bulge	149/463 (32%)	106/157 (68%)	0.31	42%	86%
Physical restraint	219/579 (38%)	36/41 (88%)	0.25	14%	99%
Cries	216/573 (38%)	39/47 (83%)	0.24	15%	98%
Spits out	229/590 (39%)	26/30 (87%)	0.21	10%	99%
Vomits	248/613 (40%)	7/7 (100%)	0.13	3%	100%

Table 3. Patient reported taste scores by medicine

	Hedonic Score		VAS Score		Tastes OK?	Composite Outcome
Drug	***Mean***	***% unacceptable***	***Mean (mm)***	***% unacceptable***	**(% "No")**	***% unacceptable***
Clarithromycin (n=26)	3.5	62%	70	58%	65%	77%
Prednisolone (n=86)	3.6	58%	68	54%	61%	70%
Amoxicillin (n=30)	2.7	37%	37	23%	33%	43%
Co-Amoxiclav (n=50)	2.5	30%	39	24%	29%	33%
Paracetamol (n=193)	2.4	22%	37	21%	21%	29%
Ibuprofen (n=92)	2.1	14%	27	12%	12%	20%

Table 4. Percentage unable to take medicine as intended

Medicine	**% spat out/vomited the medicine**	**Composite Outcome *% unacceptable***
Clarithromycin (n= 26)	23%	77%
Amoxicillin (n=30)	13%	43%
Prednisolone (n= 86)	9%	70%
Co-amoxiclav (n= 50)	6%	33%
Ibuprofen (n= 92)	2%	20%
Paracetamol (n= 193)	1%	29%

References

1. EMA. Guideline on pharmaceutical development of medicines for paediatric use EMA/CHMP/QWP/805880/2012. London. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/07/WC500147002.pdf. Accessed 18/06/2017.

2. Mistry P, Batchelor H. Evidence of acceptability of oral paediatric medicines: A review. J Pharm Pharmacol. 2016. doi:10.1111/jphp.12610.

3. Venables R, Batchelor H, Hodson J, Stirling H, Marriott J. Determination of formulation factors that affect oral medicines acceptability in a domiciliary paediatric population. Int J Pharm. 2015;480(1-2):55-62. doi:10.1016/j.ijpharm.2015.01.023.

4. Davies EH, Tuleu C. Medicines for Children: A Matter of Taste. Journal of Pediatrics. 2008;153(5):599-604. doi:10.1016/j.jpeds.2008.06.030.

5. Mistry P, Batchelor H. Methodology Used to Assess Acceptability of Oral Pediatric Medicines: A Systematic Literature Search and Narrative Review. Pediatric Drugs. 2017:1-11. doi:10.1007/s40272-017-0223-7.

6. Matza LS, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM et al. Pediatric Patient-Reported Outcome Instruments for Research to Support Medical Product Labeling: Report of the ISPOR PRO Good Research Practices for the Assessment of Children and Adolescents Task Force. Value Health. 2013;16(4):461-79. doi:http://dx.doi.org/10.1016/j.jval.2013.04.004.

7. FDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health and Quality of Life Outcomes. 2006;4(1):79. doi:10.1186/1477-7525-4-79.

8. Mistry P, Batchelor H, Tibbins C. Design of methodological tools to assess acceptability of paediatric medicines. Int J Pharm. 2016;511(2):1143-4. doi:http://dx.doi.org/10.1016/j.ijpharm.2016.06.098.

9. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A Systematic Review of Faces Scales for the Self-report of Pain Intensity in Children. Pediatrics. 2010;126(5):e1168-e98. doi:10.1542/peds.2010-1609.

10. Peryam DR, Pilgrim FJ. Hedonic scale method of measuring food preferences. Food Technology. 1957;11:9-14.

11. Katz ER, Kellerman J, Siegel SE. Behavioral distress in children with cancer undergoing medical procedures: developmental considerations. J Consult Clin Psychol. 1980;48(3):356-65.

12. Blount RL, Loiselle KA. Behavioural assessment of pediatric pain. Pain Research & Management : The Journal of the Canadian Pain Society. 2009;14(1):47-52.

13. Forestell CA, Mennella JA. More than Just a Pretty Face: The Relationship between Infant's Temperament, Food Acceptance, and Mothers' Perceptions of their Enjoyment of Food. Appetite. 2012;58(3):1136-42. doi:10.1016/j.appet.2012.03.005.

14. Motte J, Pedespan JM, Sevestre M, Chiron C. Acceptabilité et tolérance du valproate de sodium, granules à libération prolongée, en monothérapie chez l'enfant épileptique à partir de trois ans. Archives de Pédiatrie. 2005;12(10):1533-9. doi:http://dx.doi.org/10.1016/j.arcped.2005.07.009.

15. Mulla H, Buck H, Price L, Parry A, Bell G, Skinner R. 'Acceptability' of a new oral suspension formulation of mercaptopurine in children with acute lymphoblastic leukaemia.

Journal of oncology pharmacy practice : official publication of the International Society of Oncology Pharmacy Practitioners. 2015. doi:10.1177/1078155215577808.
16. Schwartz RH. Enhancing children's satisfaction with antibiotic therapy: A taste study of several antibiotic suspensions. Current Therapeutic Research. 2000;61(8):570-81. doi:10.1016/s0011-393x(00)80039-9.
17. Thompson A, Reader S, Field E, Shephard A. Open-label taste-testing study to evaluate the acceptability of both strawberry-flavored and orange-flavored amylmetacresol/2,4-dichlorobenzyl alcohol throat lozenges in healthy children. Drugs in R and D. 2013;13(2):101-7.
18. Sjövall J, Fogh A, Huitfeldt B, Karlsson G, Nylén O. Methods for evaluating the taste of paediatric formulations in children: A comparison between the facial hedonic method and the patients' own spontaneous verbal judgement. European Journal of Pediatrics. 1984;141(4):243-7.
19. Reader S, Shaw H, Hails S. taste-testing study in healthy volunteers (children) to investigate children's preference for ibuprofen or placebo suspension. 2006;7(2).
20. Smith CJ, Sammons HM, Fakis A, Conroy S. A prospective study to assess the palatability of analgesic medicines in children. Journal of Advanced Nursing. 2013;69(3):655-63.
21. Angelilli ML, Toscani M, Matsui DM, Rieder MJ. Palatability of oral antibiotics among children in an urban primary care center. Arch Pediatr Adolesc Med. 2000;154(3):267-70.
22. Gooch WM, Philips A, Rhoades R, Rosenberg R, Schaten R, Starobin S. Comparison of the efficacy, safety and acceptability of cefixime and amoxicillin/clavulanate in acute otitis media. The Pediatric Infectious Disease Journal. 1997;16(2):S21-S4.
23. Matsui D, Lim R, Tschen T, Rieder MJ. Assessment of the palatability of β-lactamase-resistant antibiotics in children. Archives of Pediatrics and Adolescent Medicine. 1997;151(6):599-602.
24. Powers JL, Gooch Iii WM, Oddo LP. Comparison of the palatability of the oral suspension of cefdinir vs. amoxicillin/clavulanate potassium, cefprozil and azithromycin in pediatric patients. Pediatric Infectious Disease Journal. 2000;19(12 SUPPL.):S174-S80.
25. Toscani M, Drehobl M, Freed J, Stool S. A multicenter, randomized, comparative assessment in healthy pediatric volunteers of the palatability of oral antibiotics effective in the therapy of otitis media. Current Therapeutic Research - Clinical and Experimental. 2000;61(5):278-85.
26. Ubaka CM, Udeogaranya OP, Ezeugwu N. Palatability of oral paediatric antibiotics commonly prescribed in a Nigerian specialist hospital. European Journal of Hospital Pharmacy: Science and Practice. 2013;20(2):122-4.
27. Dagan R, Shvartzman P, Liss Z. Variation in acceptance of common oral antibiotic suspensions. Pediatr Infect Dis J. 1994;13(8):686-90.
28. Hames H, Seabrook JA, Matsui D, Rieder MJ, Joubert GI. A PALATABILITY STUDY OF A FLAVORED DEXAMETHASONE PREPARATION VERSUS PREDNISOLONE LIQUID IN CHILDREN WITH ASTHMA EXACERBATION IN A PEDIATRIC EMERGENCY DEPARTMENT. Canadian Journal of Clinical Pharmacology. 2008;15(1):e95-e8.
29. Isa JM, Wong GK, Teraoka SS, Sera MJ, Tsushima MM, Yamamoto LG. Parental pediatric corticosteroid preferences. The American Journal of Emergency Medicine. 2001;19(1):29-31. doi:http://dx.doi.org/10.1053/ajem.2001.20025.
30. Kim MK, Yen K, Redman RL, Nelson TJ, Brandos J, Hennes HM. Vomiting of liquid corticosteroids in children with asthma. Pediatric emergency care. 2006;22(6):397-401. doi:10.1097/01.pec.0000221338.44798.6a.

31. Powers JL. Properties of azithromycin that enhance the potential for compliance in children with upper respiratory tract infections. The Pediatric Infectious Disease Journal. 1996;15(9):30-7.

32. Stafford L, Hope ME, Janney EP, Ailakis JG. Comparison of pediatric steroid mixtures. Australian Journal of Hospital Pharmacy. 1998;28(4):246-9.

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

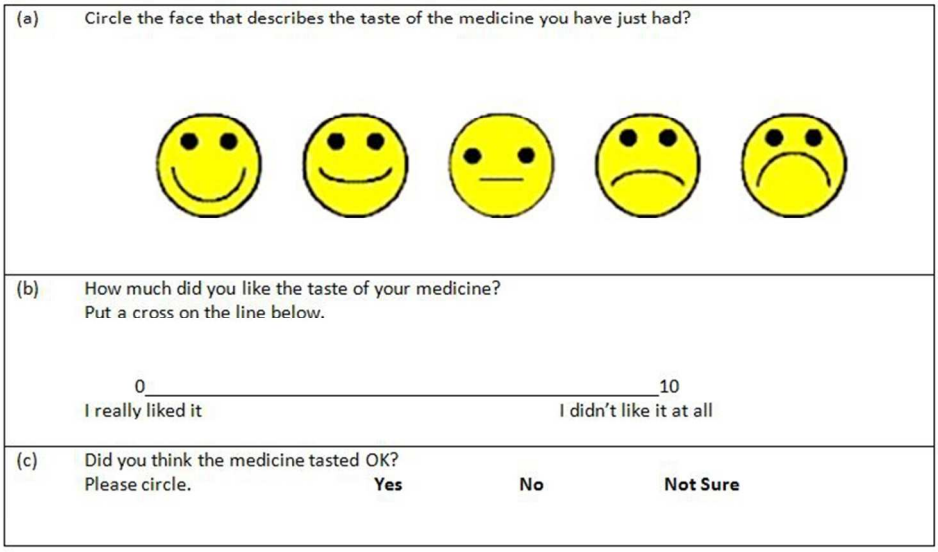


Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

67x53mm (300 x 300 DPI)

Facial expressions observed

Expression	Tick if observed prior to administration	Tick if observed during administration
Eyes squeezed shut or towards shut		
Brow bulge/lower (frown)		
Nose wrinkle		
Pursed lips		

Behaviours observed

Behaviour	Tick if observed prior to administration	Tick if observed during administration	
Child refuses medicine			=unacceptable
Child cries/screams			=unacceptable
Child requires physical restraint			=unacceptable
Child voices resistance			=unacceptable
	Tick if observed immediately after administration		
Child voices disgust			=unacceptable
Child vomits			=unacceptable
Child spits out medicine			=unacceptable
Child cries			=unacceptable

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

65x41mm (300 x 300 DPI)

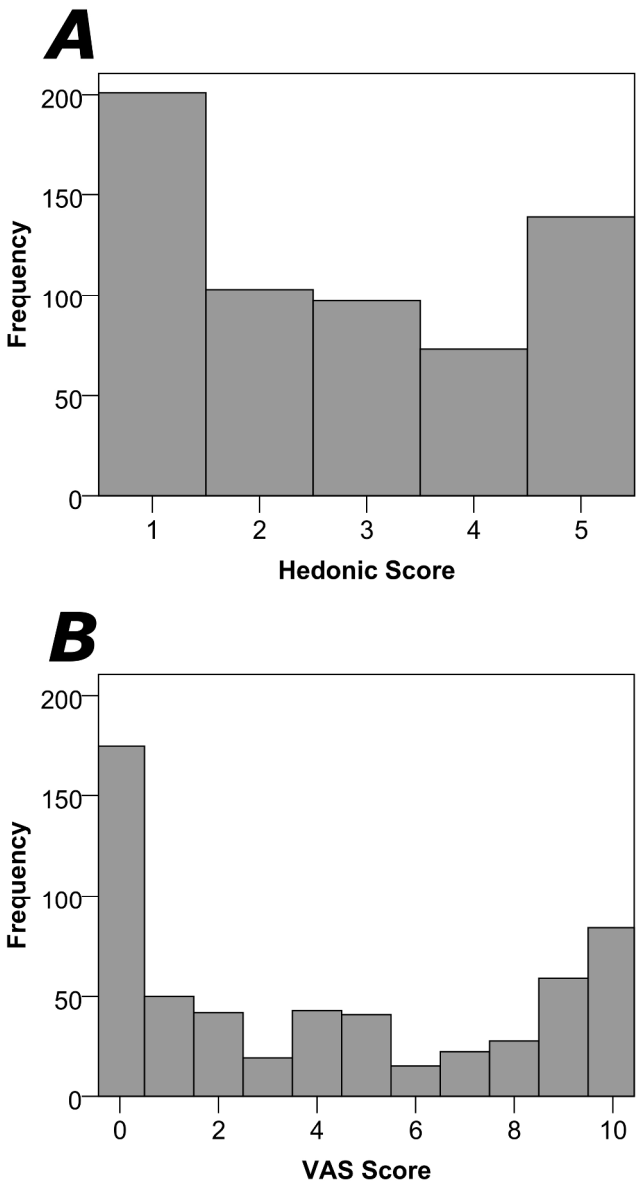
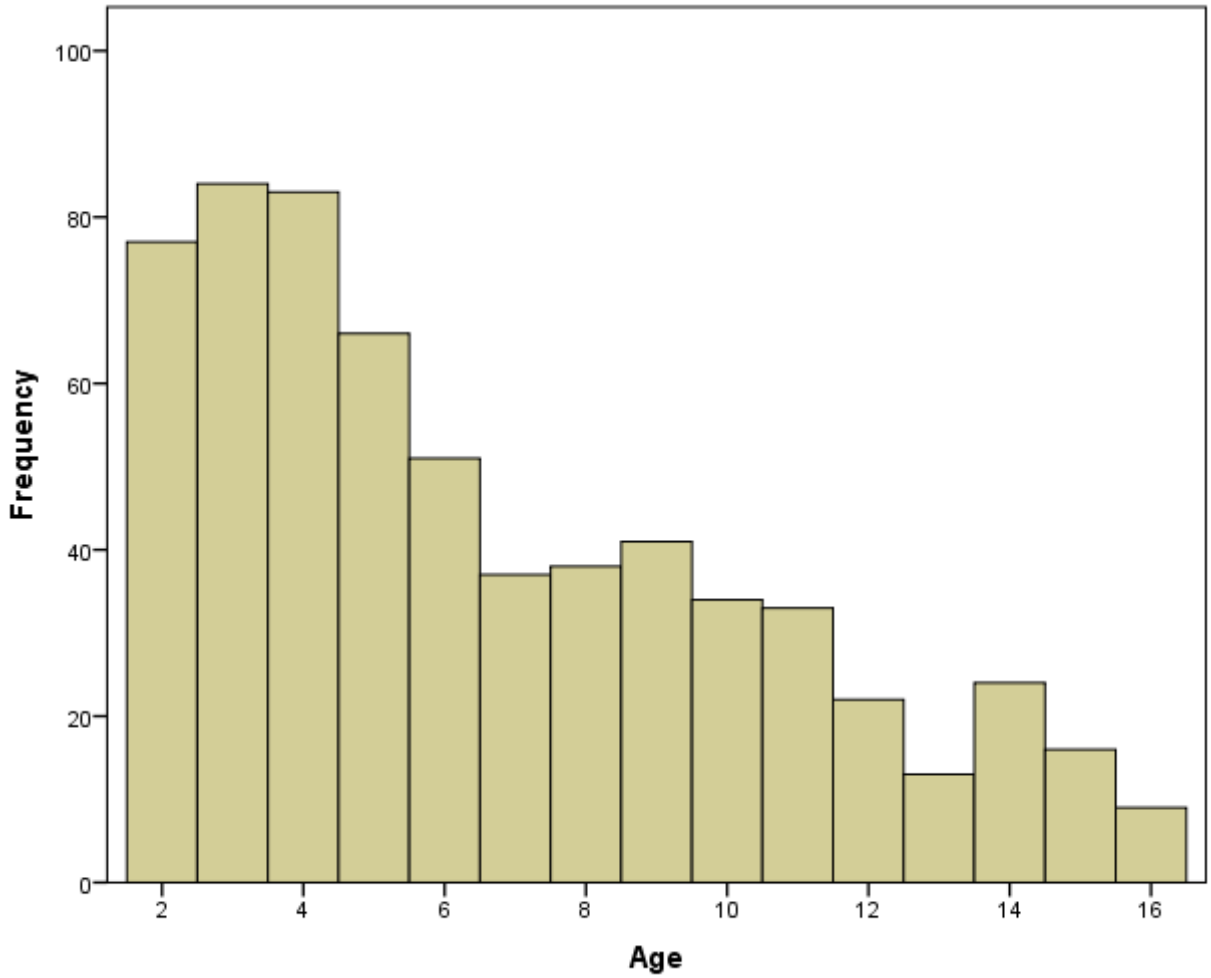
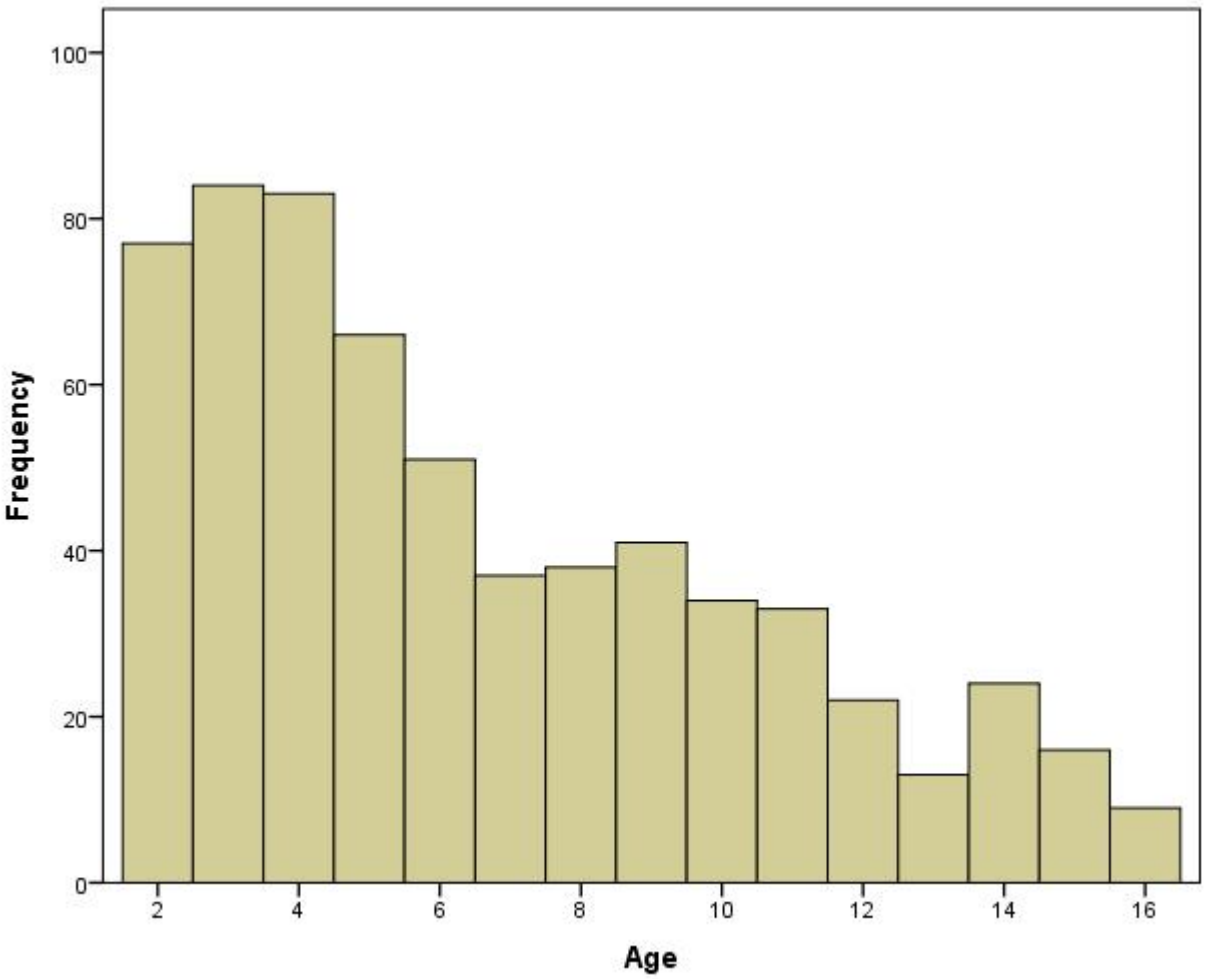


Figure 3. Hedonic and VAS score distribution

168x300mm (300 x 300 DPI)

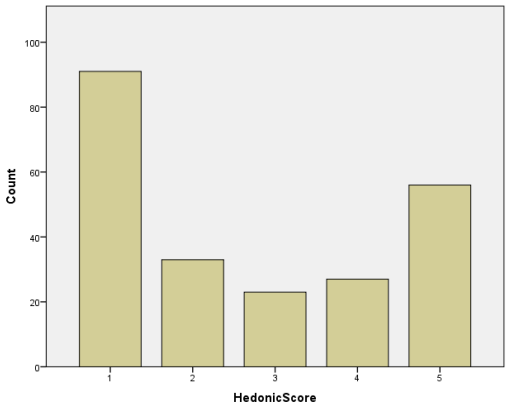
Supplementary Material 1. Distribution of participant age

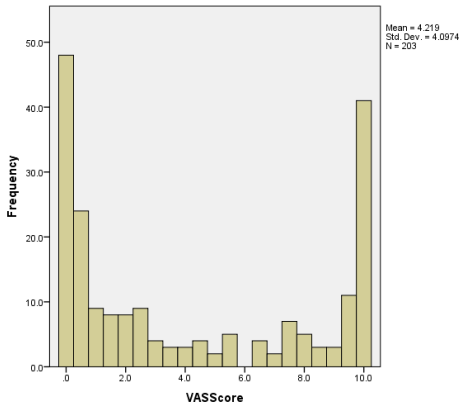




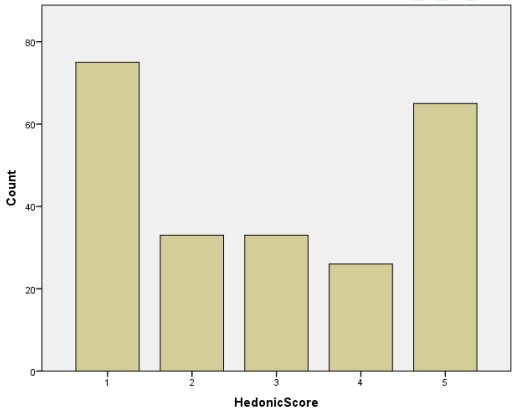
Supplementary Material 2. Age related distribution of responses from patient-reported assessment scales

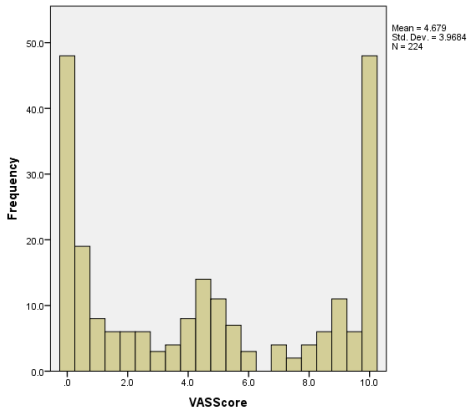
(a) Aged 2-4 years



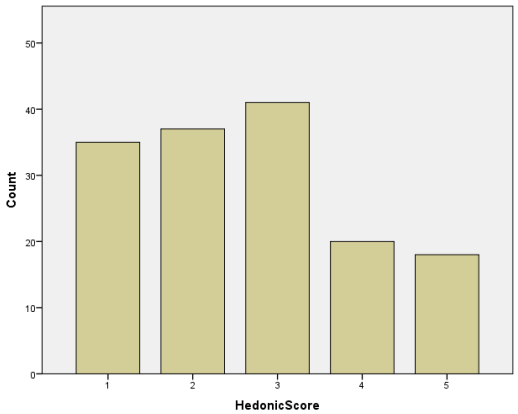


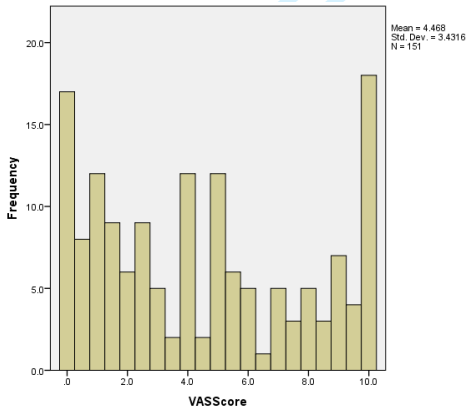
(b) Aged 5-9 years





(c) Aged 10-16 years





Supplementary material 3. Table of taste scores for the six most commonly administered drugs by brand

Drug/Manufacturer	N	Mean Hedonic Score	Mean VAS Score	Tastes OK (% "No")
Amoxicillin				
Athlone Laboratories	21	2.8	4.4	33%
Clamo	2	5.0	0.3	100%
Kent Pharmaceuticals	2	1.0	1.3	0%
Medreich PLC	1	3.0	10.0	100%
PL Holder	4	2.0	1.3	0%
Clarithromycin				
Abbott	11	3.7	7.4	82%
Klaricid	1	5.0	9.9	100%
MA Holder	1	2.0	4.2	0%
Sandoz Ltd.	13	3.4	6.7	54%
Co Amoxiclav				
Medreich PLC	1	3.0	5.0	0%
Rosemont Pharmaceutical Ltd	1	2.0	2.4	0%
Sandoz Ltd.	48	2.5	4.0	30%
Ibuprofen				
Fenpaed	9	2.1	3.4	22%
MA Holder	1	5.0	5.2	0%
McNeil Products	1	5.0	10.0	100%
Pinewood	78	2.0	2.5	10%
Reckitt Benckiser Healthcare	3	2.3	3.7	0%
Paracetamol				
Bristol Laboratories	1	3.0	4.0	100%
MA Holder	38	1.9	2.4	3%
McNeil Products	6	2.3	2.5	17%
Pinewood	113	2.6	4.3	28%
Rosemont Pharmaceutical Ltd	35	2.4	3.0	17%
Prednisolone				
Actavis	12	3.1	4.7	42%
Amdipharm	64	3.6	7.0	63%
Genetic SPA	1	3.0	5.2	0%
Teva UK Ltd	9	4.5	8.6	86%

STROBE Statement—checklist of items that should be included in reports of observational studies

This checklist was used and all items are provided within the manuscript: *Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines*

	Item No	Recommendation	Appears in manuscript (line number)
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	34-35
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Page 2 – Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	115-123
Objectives	3	State specific objectives, including any prespecified hypotheses	139-142
Methods			
Study design	4	Present key elements of study design early in the paper	145-150
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	167-180
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	168-171
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	183-208
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*172-174; 183-208*
Bias	9	Describe any efforts to address potential sources of bias	211-212
Study size	10	Explain how the study size was arrived at	216-218
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings	193-196; 219-232

		were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	220-222; 215-232
		(*b*) Describe any methods used to examine subgroups and interactions	240-241
		(*c*) Explain how missing data were addressed	227-228
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	222-224
		(*e*) Describe any sensitivity analyses	284-290

Continued on next page

Results			**Appears in manuscript**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	250-259
		(b) Give reasons for non-participation at each stage	250-259
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	237-239
		(b) Indicate number of participants with missing data for each variable of interest	250-259
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*250-259*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	261-364
		(*b*) Report category boundaries when continuous variables were categorized	301-313
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	343-364
Discussion			
Key results	18	Summarise key results with reference to study objectives	366-422
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	56-69
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	425-438
Generalisability	21	Discuss the generalisability (external validity) of the study results	433-438
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	90-95

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines in an inpatient paediatric population

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-021961.R2
Article Type:	Research
Date Submitted by the Author:	02-May-2018
Complete List of Authors:	Mistry, Punam; University of Birmingham, School of Pharmacy Stirling, Heather; University Hospitals Coventry and Warwickshire NHS Trust, Paediatrics Callens, Claire; NIHR Clinical Research Network:West Midlands – Young Person's Steering Group (YPSG), CRN: Children Hodson, James; Queen Elizabeth Hospital Birmingham, Institute of Translational Medicine Batchelor, Hannah; University of Birmingham, School of Pharmacy
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Patient-centred medicine, Pharmacology and therapeutics
Keywords:	THERAPEUTICS, PAEDIATRICS, palatability, ORAL MEDICINE

Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines in an inpatient paediatric population

Punam Mistry[1], Heather Stirling[2], Claire Callens[3], James Hodson[4] and Hannah Batchelor[1] on behalf of SPaeDD-UK project

(Smart Paediatric Drug Development – UK, accelerating paediatric formulation development http://www.paediatricscienceuk.com)

[1] School of Pharmacy, Institute of Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT, UK

[2] University Hospital Coventry and Warwickshire NHS Trust, Clifford Bridge Road, Coventry, CV2 2DX, UK

[3] NIHR Clinical Research Network West Midlands, Institute of Research and Development, Birmingham Research Park, Vincent Drive, Birmingham, B15 2SQ, UK

[4] Institute of Translational Medicine, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston, Birmingham, B15 2GW, UK

Affiliation: Pharmacy and Therapeutics, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT, United Kingdom

Corresponding author: Hannah Batchelor, School of Pharmacy, Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, B15 2TT
Email: h.k.batchelor@bham.ac.uk
Phone: +44 (0)121 414 3717

Abstract

Objective: To evaluate the age-appropriateness and suitability of patient reported outcome measures to assess the acceptability of the taste of oral liquid medicines in children.

Design and setting: An observational mixed methods study involving children aged 2-16 years old taking oral liquid medicine in paediatric inpatient wards across the West Midlands (UK). Assessment tools included: patient-reported scores on the taste of medicines via a 5-point facial hedonic scale; a visual analogue scale (VAS); a question, "Did you think the medicine tasted OK?" and researcher observations of facial expressions and behaviours immediately before, during and after administration.

Results: 611 children participated. The percent unable to complete the scales were 7% (n=46) for the VAS; 2% (n=15) for the hedonic scale and 1% (n=7) for the question about taste. Significant correlations (Spearman's rho) were observed between the patient reported outcome measures: 0.80 and 0.78 for the taste question and hedonic and VAS respectively and 0.84 for the hedonic and VAS. Researcher observations demonstrated the ability of the patient to take the medicine as intended but did not provide sensitive measures of taste. 5% of administrations were not taken as intended by the children. Medicines known to have poor taste (clarithromycin and prednisolone) showed mean hedonic and VAS scores of ≥3.5 and >65mm respectively.

Conclusions: Patient reported outcome measures correlate with each other and are a useful means to assess the taste (and acceptability) of medicines. Hedonic scales are better understood by children and should be the first choice tool in the assessment of medicines taste.

Key words: medicines, taste, acceptability, age-appropriate, hedonic scale, paediatric medicine, patient-reported outcome measures, VAS

Strengths and limitations of this study

- This is the first study to compare methodologies to assess the acceptability of taste of liquid medicines in a large UK based paediatric population aged 2-16 years
- The sample size in this study provided large data sets of six key medicines, which provides a representative comparison for newly developed products
- This study was conducted within an inpatient environment and the acceptance of taste of medicines in a domiciliary environment may differ
- The study design captured the most relevant aspects of acceptability of taste whilst minimising the burden to participants, it was not possible to measure every aspect. Suggestions for future research include measurement of: impact of the devices used to administer medicines; use of alternative hedonic scales (e.g. those with 2-9 faces); alternative anchor phrases; an endpoint of neutral and not positive within a hedonic scale; further exploration of medicines that tasted OK as well as those with a reported negative taste.

What is already known about this subject?

- New medicines for children must be demonstrated to be acceptable to a paediatric population
- Measurement of acceptability of the taste of medicines is complex and a wide range of methods have been used previously, making comparison between studies complex.
- There is a need for age-appropriate reproducible and reliable methods to measure the acceptability of medicines

What this study adds

- Patient reported outcome measures offer a pragmatic means to assess the taste of medicines in children aged from 2-16 years

- 5-point hedonic scales were better understood compared to visual analogue scales in children aged 2-16 years
- Although 41% of medicines were reported to have unacceptable taste only 5% were so bad that they could not be taken as intended

Funding Statement

This work was conducted as part of the SPaeDD-UK project (Smart Paediatric Drug Development – UK, a project co-funded by Innovate UK and the contributing companies of AstraZeneca, Bristol Myers Squibb, GlaxoSmithKline, Juniper Pharmaceuticals and Pfizer. (http://www.paediatricscienceuk.com).

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Data sharing statement

Additional data is available in the Supplementary files. The full data set is held by the corresponding author, please email with any requests for extra data.

Authorship contributions

Punam Mistry contributed to the acquisition, analysis and interpretation of the data.

Heather Stirling contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Claire Callens contributed to the design of the study and acquisition of the data

James Hodson contributed to the design, statistical analysis and interpretation of the data and revising the manuscript following reviewers' comments.

Hannah Batchelor contributed to the design, acquisition, analysis and interpretation of the data and revising the manuscript following reviewers' comments. She is the corresponding author for this work

Word count of manuscript body: 3836

1. Introduction

Measurement of acceptability of paediatric medicines is important for two major reasons: (i) poor acceptability is likely to be associated with poor compliance and therefore suboptimal therapy; (ii) regulatory guidelines dictate that paediatric medicines' acceptability should be assessed (preferably in children) [1]. Patient acceptability has been defined as, "an overall ability of the patient and caregiver (defined as 'user') to use a medicinal product as intended (or authorised)" [1]. A major barrier in the assessment of acceptability and subsequent interpretation of results is the lack of guidance on how an acceptability test should be conducted and the criteria used to determine acceptability in children [2]. It is unlikely that a single acceptability test will be valid for all medicines, due to the need to consider the benefit:risk balance on an individual basis.

Taste was previously identified as the biggest barrier in adherence and compliance in a paediatric population [3]. Previous techniques to evaluate acceptability focussed on taste and were based on methodology adapted from sensory analysis undertaken in the food industry. A variety of scales have been used to measure the acceptability of medicines, yet it remains unclear whether any of the scales is better for a particular purpose with regard to validity, reliability, feasibility, and preference [4, 5]. Patient reported measures of taste offer a pragmatic means to collect data on products. However, there is often concern about the reliability of these measures, particularly in a paediatric population [6]. Sensory analysis is commonly used to show a preference for one product over another, but this is not always relevant to medicines as often there is no alternative available. In addition, it can be rationalized that the therapeutic benefits of medicines outweigh the negative taste. It is complex to distinguish between acceptability and likeability of a medicine's taste, yet this difference may be critical for appropriate interpretation of resulting data.

The validity of any new method relies on how well it compares to other measures [7]. The development of a reliable method to assess the taste of medicines requires comparison of a variety of tools. This study compares a range of methods to assess the taste of medicines within a large

paediatric inpatient population. Independent researcher observations are also compared to the self-reported data. The results of this study will be used to propose a suitable method that can be used for future taste assessments.

2. Materials and Methods

Three patient-reported outcome (PRO) measures were compared to each other, and to researcher observations of medicines administration in an observational mixed methods study.

2.1. Patient and Public Involvement

Bespoke PRO tools were developed for this study based on previous methodologies and in consultation with the National Institute for Health Research (NIHR) Children Specialty's Young Person's Advisory Group (West Midlands) [8]. The young people (aged 11-18 years) reviewed the tools and provided feedback that the tools were age-appropriate. The same young people provided feedback on the trial materials including information sheets and how to minimise the burden to participants during the conduct of the study. The results are available to participants as a poster summary from the corresponding author's personal webpage (www.hannahbatchelor.com); this poster was also reviewed by the young person's group.

2.2. Patient reported outcome measures used

The hedonic scale selected was a genderless image where the mouth was the only expressive facial feature (see Figure 1(a)). The images were yellow to ensure they stood out from the background paper. They were obtained from S-cool the revision website (http://www.s-cool.co.uk/gcse/food-technology/systems-and-control/revise-it/sensory-evaluation (accessed December 2015)). Children and young people preferred simple faces and felt that this would be most appropriate for the youngest age group [8]; anchor phrases were not included on the hedonic scale to keep this tool basic. The visual analogue scale (VAS) included anchor phrases selected by the young persons'

group as the most clear and relevant [8], these were used at the extreme ends of the continuous scale.

The direction of change was from positive to negative, which corresponds to the extensive data on hedonic scales used for pain assessment in children [9]. Previous studies in adult studies of foods showed no difference based on structural variations that read from positive to negative or vice versa [10].

The third PRO was a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure.

2.3. Participants and Setting

Paediatric inpatients aged between 2 to 16 years taking any oral liquid medicine as part of their medical care were recruited using convenience sampling, from inpatient wards at 11 sites across the West Midlands. Informed consent was obtained from the parent or legal guardian of the participating child and for children over 12 years of age assent was also required. The study was conducted between December 2015- April 2016. Ethical approval was granted by London Surrey borders ethical committee (REC reference: 15/LO/1253; IRAS project ID: 179684).

Demographic information was obtained on participant's age and whether this was their first dose of the medicine under evaluation. The medicine name; brand; manufacturer; concentration (strength); dose administered and product batch number was recorded.

Each participant was observed by a researcher as they took their medicine. Some medicines were provided to the patient as an oral liquid following extemporaneous preparation within the clinical setting, for example, dispersion of prednisolone tablets immediately prior to dosing. Depending on the capability of the child, the medicines were either self-administered, or administered by nursing staff and/or parents. Participants were asked not to mix the medicine with any other food product, as this might influence the participant's responses.

2.4. Patient-reported outcome tools

Participants were asked to complete (a) 5-point facial hedonic scale and (b) 10cm VAS (Figures 1(a) and (b)) immediately after administration of their medicine; both scales were provided on separate paper documents in a randomized order. Children were free to ask for support in completing the questionnaires from parents, nursing staff or the researcher present. The cognitive function of children was not assessed and age may not always predict a child's ability to complete the questionnaire, therefore all children were free to ask for support if requried. Both reporting documents included a third PRO (Figure 1 (c)) as a question, 'Did you think the medicine tasted OK?' with the response options of: yes, no, not sure. The purpose behind this question was to endorse the reliability of the participant's reporting from both scale-based questionnaires.

Figure 1

The data from the scales (Figure 1 (a) and (b)) were transformed into numbers for subsequent statistical analysis. Hedonic expressions were converted to scores of 1 to 5 with 1 being the most positive expression and 5 the most negative. VAS scores were reported as measurements from the extreme left hand side of the scale in mm; this provided scores of 0-100 mm (although scale was written as 0-10 cm).

2.5. Researcher Observations

A facial expression and behavioural scale to capture researcher observations was a 12-point tick chart designed in-house (Figure 2), based on an existing scale used to measure behavioural distress in children undergoing medical procedures [11], the Procedural Behavior Rating Scale (PBRS). The PBRS was developed specifically for infants and children with cancer, who were undergoing bone marrow aspirations or lumbar punctures. It contained 13 behavioural codes including: cry, cling, pain verbal, flail, stall, scream, refuse position, restrain, emotional support, muscular rigidity and requests termination [11]. The in-house tool was based on this scale plus other scales used to measure

behavioural distress in children [12]. The facial expressions included on the scale were derived from previous studies that assessed food-liking in children based on their facial expressions; typically negative tastes are associated with eyes squeezing, brow bulge, nose wrinkle and pursed lips [13].

Figure 2

A series of nine short films or still pictures of children were made available to researchers participating in the study to assess the inter-rater agreement in the facial expressions displayed.

2.6. Statistical analysis

A sample size was not fixed for this study at the outset as there was no appropriate power calculation. Advice from a statistician was to ensure we had sufficient numbers (<50) paired data to ensure that selectivity and sensitivity of the methods could be demonstrated.

Initially, the mean age of participants that were unable to understand the assessments were compared to those who could using Mann-Whitney tests. The same approach was also used to compare the scores for patients receiving their first dose, relative to those who had previously received the medicine. Age was then divided into categories, and the proportions of participants scoring in the extreme categories for the scales were compared using Fisher's exact tests.

Spearman's correlation coefficients were used to assess the degree of correlation between the assessments for the cohort as a whole, and within each of the age categories. In this analysis, the "Did you think the medicine tasted OK?" question was treated as an ordinal scale, with the response of Don't Know being between Yes and No. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available

The assessments were then dichotomised, and compared using McNemar's tests to assess for marginal homogeneity in the 2x2 tables, and Kappa to assess agreement. The three assessments were then combined into a composite score, which was compared with reported behaviours using

Kendall's tau correlation coefficients. All analyses were performed using IBM SPSS 22 (IBM Corp. Armonk, NY), and $p<0.05$ deemed to be indicative of statistical significance throughout.

3. Results and discussion

Data were available for 628 administrations to 611 children aged between 2-16 years. The median participant age was 6 years. Further details on the distribution of the participant ages can be found in supplementary material 1.

To ease analysis of data, the population for this study was stratified by age into three groups: 2-4 years (n=237); 5-9 years (n=227) and 10-16 years (n=147).

The medicine was administered as the first dose in 162 cases. There was no evidence of a significant difference in the hedonic or VAS scores between those receiving their first dose of a medicine, compared to those who had previous administrations ($p=0.336$, 0.909 respectively). For all subsequent analysis the data was pooled for those receiving their first and subsequent doses of medicine.

3.1. Completeness of patient-reported assessment scales

The assessment scales were not completed by all of the study participants. The VAS had the lowest completion rate, where 46 (7%) were not completed due to lack of understanding by the child, compared to 15 (2%) for the hedonic scale. There were 7 (1%) cases where the child did not understand the question, "Did you think the medicine taste OK?". The range and mean age of those unable to complete the scales were 2-5 years with a mean of 2.4 years for the hedonic; 2-6 years with a mean of 3.1 years for the VAS and 2-3 years with a mean of 2.3 years for the taste question. In each case, participants unable understand the assessment methods were significantly younger than the remainder of the cohort ($p<0.001$ each assessment). The cognitive function of children was not

assessed within this study and there was an assumption of cognitive normal for age for all participants.

3.2. Distribution of responses to patient-reported assessment scales

In the patient reported scales (VAS and hedonic), the responses tended to be biased towards the extreme responses, with 56% of responses being in the highest or lowest categories for hedonic score (Figure 3a) and, 55% within 1cm from either end of the scale in the VAS (Figure 3b).

Figure 3

The use of the extreme ends of the scales was greater in the younger populations ($p<0.001$) (data is shown in Supplementary material 2). For those that completed the hedonic scale, 64% of those aged 2-4 years used the extreme ends of the scale compared to 60% of those aged 5-9; 35% of those aged 10-16 years. A similar trend was observed in the VAS data where 66% of those aged 2-4 years were within 10mm of either end of the scale compared to 60% of 5-9 year olds and 39% of 10-16 year olds.

3.3. Reliability of assessment, "Did you think the medicine tasted OK?"

The question, "Did you think the medicine tasted OK?" was asked to participants on two occasions for each medicine administration; once following the hedonic and once following the VAS (which were presented in a randomized order). In the 600 cases where participants responded to both questions, the same response was given in 569 (95%) of cases, giving a Kappa statistic of 0.91. In cases where the answers were not consistent, the median age of the respondents was 5 years (mean age = 6.3 years). For subsequent analysis, the two versions of this question were merged, with discrepancies resolved by taking the most positive response given by the participant.

3.4. Correlation between patient reported outcome measures

Significant correlations were observed between the hedonic scale score, VAS and "Did you think the medicines tasted OK?" question (all p<0.001 Table 1), with the strongest correlation observed between the hedonic and VAS scores (Spearman's rho=0.84). The weakest correlations were consistently observed in the youngest patients (age 2-4 years), implying that this group of patients had the lowest consistency in scores given across the different assessments. However, despite this, the consistency between the scores was still reasonable, with correlation coefficients ranging from 0.68- 0.77.

The assessments were dichotomised, to identify those responses that classified a medicine as tasting acceptable. For the hedonic scale, this was defined as a neutral to positive response (the leftmost three faces of Figure 1a). For the *"Did you think the medicine tasted OK?"* question, grouping the "yes" and "not sure" responses was considered most appropriate, as the overall purpose is the assessment of acceptability of taste of a medicine; the response "no" is clearly aligned to an unacceptable taste, making other responses acceptable. The level of agreement between the resulting dichotomised scores was high (90%), with a Kappa statistic of 0.782, and no indication of bias (McNemar's test: p=0.519).

Correlation of the hedonic scale to the "tastes OK" question is simple; however interpretation of the VAS is more complex. The VAS had anchor phrases at either end to maintain the continuous scale, so it was not obvious where neutral or positive was ranked on the scale. Initially, an arbitrary cut-off of <50mm in the VAS was used to define acceptability, which gave reasonable agreement with both the dichotomised hedonic (88%) and "tastes OK" (86%) measures. However, McNemar's test indicted significant bias in both cases (p<0.001), with a tendency for the VAS score of 0-50mm to underestimate the number of tastes-acceptable responses, compared to the hedonic and "tastes OK" assessments. As a result, alternative VAS cut-offs were explored, with 0-70mm selected to approximately match the cut-off on the hedonic scale where the neutral face becomes the negative.

Using 70mm rather than 50mm in the VAS marginally increased the agreement with both the dichotomised hedonic and "tastes OK" measures to 91% but, more importantly, eliminated the previously observed bias (McNemar's test p=1.000, 0.683 respectively). All subsequent analysis used the cut-off of >70mm as a measure of unacceptable taste.

3.5. Associations with researcher observations and patient reported outcome measures

One patient did not have a record of facial expression/behaviours, hence they were excluded, and the analysis was based on n=620 cases. Associations between facial expressions and behaviours (listed in Figure 2) were linked to cases where the medicines were reported to have an unacceptable taste, based on a composite outcome (Table 2). This was defined using the criteria defined previously. Where responses were not recorded for all criteria, the outcome was calculated based on those that were available. In total, 255/620 (41%) of assessments identified the taste of medicines to be unacceptable. The associations between this outcome, and the various facial expressions and behaviours that were recorded were then assessed.

Table 2 shows the incidence of the facial expression or behaviour as a fraction of the total population. The behaviours are listed in order of Kendall's tau correlation coefficients, with those behaviours most strongly associated with unacceptable taste having the highest value of tau. Based on this metric, voicing disgust was the behaviour most strongly associated with negative taste. Of the 105 patients that voiced disgust, 95 (90%) reported the taste of the medicine to be negative. This rate was almost three-fold higher than the 160/515 (31%) who reported negative taste after not voicing disgust. Ordering the data in this way puts 'vomits' in last place, despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable. Since so few children vomited (n=7), the proportion of the total number of children who identified the taste as unacceptable and also displayed this behaviour (i.e. the sensitivity) was low, at 3%. Hence, although a vomiting child is almost certain to find the taste unacceptable, using this behaviour in isolation as a predictor of negative taste would miss the vast majority of patients who reported taste to be negative

Inter-rater agreement assessed via the use of short films and images were mixed; prevalent expressions were detected in >95% of cases, whereas some mild expressions were only detected in 40-50% of those viewing the images.

3.6. Analysis of medicine-specific taste assessment

Fifty-seven different drugs were observed in this study and the six most commonly administered were paracetamol, ibuprofen, prednisolone, co-amoxiclav, amoxicillin and clarithromycin, which made up 76% (n=477) of the total data set.

Comparisons were made across the six most commonly prescribed drugs. Table 3 ranks these drugs in order of taste, and reports the means scores on the hedonic scale and VAS, as well as the proportion of patients answering "no" to the, "Did you think the medicine tasted OK?" question.

The drugs can be divided into three groups based on this data: clarithromycin and prednisolone were found to be the worst tasting; amoxicillin and co-amoxiclav were mid-range, and ibuprofen and paracetamol were the best tasting medicines. The effect of brand was also investigated and the data is presented in Supplementary material 3.

In addition to reports of taste, the proportion of children who refused, vomited or spat out the medicines, was calculated and classified as unable to "use a medicinal product as intended". In total, this occurred in 33/620 (5%) administrations, which are reported for the six most commonly used drugs in Table 4.

Clarithromycin was the most commonly not taken as intended and was also the drug most frequently identified as having unacceptable taste, based on the previously defined composite outcome. However, there was insufficient data to suggest that the taste of the medicine was directly related to the ability to take the medicine as intended. Children may vomit due to their underlying illness rather than as a direct result of the taste of their medicine.

4. Discussion

Few studies have categorised acceptability of the taste of medicines. The results within this study agree with previous reports, where a neutral to positive hedonic response indicates acceptable taste [14-17]; Sjovall et al (1984) compared two brands of penicillin and reported that the acceptable taste mean hedonic score was within the neutral to positive range and an unacceptable taste was in the negative range [18]. Children were free to ask for support in completing the PRO measured and we did not collect data on how many received help in this aspect; it would be of value to consider how many, particularly in the youngest age group received support. Many of the children aged 2-5 years were able to provide reliable data on the taste of medicines demonstrating that the scales and questions used within this study are suitable for very young participants.

4.1. Interpretation of facial expressions and behaviours

Voicing disgust was the behaviour most strongly associated with unacceptable taste whilst, perhaps counterintuitively, vomiting was the behaviour least strongly correlated with unacceptable taste. Despite the fact that 100% of patients who vomited found the taste of their medicine to be unacceptable, only a small number of patients vomited (n=7). As a result, whilst vomiting was a highly specific predictor of unacceptable taste, it had very low sensitivity. A similar pattern was also observed for the other facial expressions and behaviours, with sensitivity ranging from 10%-51%, and specificity from 84%-99%. This can be interpreted as indicating that, whilst it was uncommon to observe these facial expressions and behaviours in patients who found the taste of the medicine acceptable, displaying facial expressions and behaviours was not a strong indicator of unacceptability.

The behaviours used to inform the researcher observations were not always clearly defined; for example the use of physical restraint was not explicitly stated and further work is required to better understand what physical restraint may be considered acceptable.

The explicit definition of an acceptable medicine being "an overall ability of the patient and caregiver (defined as 'user') to use a medicinal product as intended (or authorised)" (Kozarewicz, 2014), includes the patient/caregiver's ability to access the medicine and comply with packaging requirements and for this study to demonstrate that the medicine was swallowed without incident. In total, 33/620 (5%) of medicines were spat out or vomited, and therefore unacceptable within this in patient population. This demonstrates that, although some of the behaviour and expressions observed may link more strongly to a negative taste, they do not automatically mean that the medicine was unacceptable.

In future studies, observations should ensure that the medicine was taken as intended; this may require a simple tool to ensure that the dose was completely swallowed without spitting out or

vomiting. There is no need to include additional observations, as these were not strongly correlated to patient reported outcomes on the taste of medicines.

4.2. Medicine-specific taste assessment

Our results correlate well with other reports in the literature on the taste of medicines. Both paracetamol and ibuprofen have previously been reported to demonstrate acceptable taste (mid-range in a 5 point hedonic scale) in paediatric populations [19, 20]. Co-amoxiclav and amoxicillin were reported to have acceptable taste to 57-67% of children within this study, which is consistent with previous literature, which has reported co-amoxiclav to be preferred to other antibiotics in terms of taste, including amoxicillin [21-27]. Previous work suggested that 60% of children accepted amoxicillin without problems and 63% of children liked the taste of co-amoxiclav [27]. Prednisolone and clarithromycin were the least liked medicines, again consistent with published data, with results showing predominantly negative scores in hedonic scales or the equivalent of >70mm in a VAS [28-30, 23, 31, 32, 3]. This study identified differences in acceptance of certain brands of medicines; however, there was insufficient sample numbers to undertake a detailed analysis of the impact of brand on taste and acceptance. A future study that explored brand preference would be of value to patients and those purchasing commonly used medicines.

Using our criteria, the proportion of patients classifying taste as acceptable for the most liked of the frequently prescribed medicines (ibuprofen and paracetamol) was over 70%; the proportion classifying taste as acceptable for the least liked medicines (prednisolone and clarithromycin) was less than 30%. This study provides some standardised values for medicines used within a UK population, the thresholds reported here provide guidance for future medicines development where in addition to taste scores the overall risk-benefit of the medicine will need to be considered.

4.3. Recommended tools to assess acceptability

This study has correlated three simple patient reported measures of medicines taste acceptability. It has also provided comparative data from existing medicines. Regulations mandate that all new medicines need to be demonstrated to be acceptable to children [1]. This study provides pragmatic and reliable tools to conduct this assessment. Furthermore, comparison of the results from a new medicine using these tools can be directly compared to existing medicines to support evidence of acceptance.

5. Conclusions

This study has generated data on the taste of medicines commonly used in paediatric populations aged 2-16 years. The results of this study suggest that patient reported outcome measures are a reliable and valid assessment of the taste of children's medicines, for children aged from 2-16 years. These assessment tools offer a mechanism to evaluate the taste of other medicines (either novel products and formulations or medicines used orally in an off-label or unlicensed manner) to generate comparative data on the taste of medicines.

The data from this study coupled with previous literature on the taste of medicines provides evidence to suggest criteria to demonstrate acceptability of taste of medicines.

Our results suggest that criteria to demonstrate acceptability of taste are: a mean VAS score of <70mm; a mean hedonic score of ≤3 (neutral or positive face) and a non-negative response to the "Tastes OK?" question. Pragmatically, there is no need to use all methods. As the hedonic scale was understood across the widest age range, this should be the first choice method going forwards.

It would be prudent to ensure that any new product exceeds these scores to demonstrate that it is likely to have acceptable taste in practice.

Acknowledgements

The authors would like to thank the NIHR Clinical Research Network: West Midlands – Young Person's Steering Group (YPSG) for their input and advice in the development, conduct and dissemination of this study.

All patients, families and researchers at the participating sites are acknowledged for their participation in this study.

Figure Legends

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

Figure 3. Hedonic and VAS score distribution

Table headings

Table 1. Spearman's rho correlation values between patient-reported outcome measures.

	"Tastes OK" vs. Hedonic	**"Tastes OK" vs. VAS**	**Hedonic vs. VAS**
Overall	***0.80***	***0.78***	***0.84***
Age (Years)			
2-4	0.77	0.68	0.76
5-9	0.83	0.85	0.90
10-16	0.80	0.81	0.86

Table 2. Relationship between facial expression or a behaviour and a patient report of an unacceptable tasting medicine

	Cases Where Taste was Reported as unacceptable (n=255)				
Behaviour	***Not Displayed***	***Displayed***	**Tau**	**Sens.**	**Spec.**
Voices disgust	160/515 (31%)	95/105 (90%)	0.45	37%	97%

Eyes squeezed	131/460 (28%)	124/160 (78%)	0.44	49%	90%
Nose Wrinkle	125/433 (29%)	130/187 (70%)	0.38	51%	84%
Voices resistance	184/539 (34%)	71/81 (88%)	0.37	28%	97%
Refusal	192/551 (35%)	63/69 (91%)	0.36	25%	98%
Pursed Lips	168/505 (33%)	87/115 (76%)	0.33	34%	92%
Cries/Screams	201/559 (36%)	54/61 (89%)	0.32	21%	98%
Brow bulge	149/463 (32%)	106/157 (68%)	0.31	42%	86%
Physical restraint	219/579 (38%)	36/41 (88%)	0.25	14%	99%
Cries	216/573 (38%)	39/47 (83%)	0.24	15%	98%
Spits out	229/590 (39%)	26/30 (87%)	0.21	10%	99%
Vomits	248/613 (40%)	7/7 (100%)	0.13	3%	100%

Table 3. Patient reported taste scores by medicine

	Hedonic Score		VAS Score		Tastes OK?	Composite Outcome
Drug	***Mean***	***% unacceptable***	***Mean (mm)***	***% unacceptable***	**(% "No")**	***% unacceptable***
Clarithromycin (n=26)	3.5	62%	70	58%	65%	77%
Prednisolone (n=86)	3.6	58%	68	54%	61%	70%
Amoxicillin (n=30)	2.7	37%	37	23%	33%	43%
Co-Amoxiclav (n=50)	2.5	30%	39	24%	29%	33%
Paracetamol (n=193)	2.4	22%	37	21%	21%	29%
Ibuprofen (n=92)	2.1	14%	27	12%	12%	20%

Table 4. Percentage unable to take medicine as intended

Medicine	**% spat out/vomited the medicine**	**Composite Outcome *% unacceptable***
Clarithromycin (n= 26)	23%	77%
Amoxicillin (n=30)	13%	43%
Prednisolone (n= 86)	9%	70%
Co-amoxiclav (n= 50)	6%	33%
Ibuprofen (n= 92)	2%	20%
Paracetamol (n= 193)	1%	29%

References

1. EMA. Guideline on pharmaceutical development of medicines for paediatric use EMA/CHMP/QWP/805880/2012. London. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/07/WC500147002.pdf. Accessed 18/06/2017.
2. Mistry P, Batchelor H. Evidence of acceptability of oral paediatric medicines: A review. J Pharm Pharmacol. 2016. doi:10.1111/jphp.12610.
3. Venables R, Batchelor H, Hodson J, Stirling H, Marriott J. Determination of formulation factors that affect oral medicines acceptability in a domiciliary paediatric population. Int J Pharm. 2015;480(1-2):55-62. doi:10.1016/j.ijpharm.2015.01.023.
4. Davies EH, Tuleu C. Medicines for Children: A Matter of Taste. Journal of Pediatrics. 2008;153(5):599-604. doi:10.1016/j.jpeds.2008.06.030.
5. Mistry P, Batchelor H. Methodology Used to Assess Acceptability of Oral Pediatric Medicines: A Systematic Literature Search and Narrative Review. Pediatric Drugs. 2017:1-11. doi:10.1007/s40272-017-0223-7.
6. Matza LS, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM et al. Pediatric Patient-Reported Outcome Instruments for Research to Support Medical Product Labeling: Report of the ISPOR PRO Good Research Practices for the Assessment of Children and Adolescents Task Force. Value Health. 2013;16(4):461-79. doi:http://dx.doi.org/10.1016/j.jval.2013.04.004.
7. FDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health and Quality of Life Outcomes. 2006;4(1):79. doi:10.1186/1477-7525-4-79.
8. Mistry P, Batchelor H, Tibbins C. Design of methodological tools to assess acceptability of paediatric medicines. Int J Pharm. 2016;511(2):1143-4. doi:http://dx.doi.org/10.1016/j.ijpharm.2016.06.098.
9. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A Systematic Review of Faces Scales for the Self-report of Pain Intensity in Children. Pediatrics. 2010;126(5):e1168-e98. doi:10.1542/peds.2010-1609.
10. Peryam DR, Pilgrim FJ. Hedonic scale method of measuring food preferences. Food Technology. 1957;11:9-14.
11. Katz ER, Kellerman J, Siegel SE. Behavioral distress in children with cancer undergoing medical procedures: developmental considerations. J Consult Clin Psychol. 1980;48(3):356-65.
12. Blount RL, Loiselle KA. Behavioural assessment of pediatric pain. Pain Research & Management : The Journal of the Canadian Pain Society. 2009;14(1):47-52.
13. Forestell CA, Mennella JA. More than Just a Pretty Face: The Relationship between Infant's Temperament, Food Acceptance, and Mothers' Perceptions of their Enjoyment of Food. Appetite. 2012;58(3):1136-42. doi:10.1016/j.appet.2012.03.005.
14. Motte J, Pedespan JM, Sevestre M, Chiron C. Acceptabilité et tolérance du valproate de sodium, granules à libération prolongée, en monothérapie chez l'enfant épileptique à partir de trois ans. Archives de Pédiatrie. 2005;12(10):1533-9. doi:http://dx.doi.org/10.1016/j.arcped.2005.07.009.
15. Mulla H, Buck H, Price L, Parry A, Bell G, Skinner R. 'Acceptability' of a new oral suspension formulation of mercaptopurine in children with acute lymphoblastic leukaemia.

Journal of oncology pharmacy practice : official publication of the International Society of Oncology Pharmacy Practitioners. 2015. doi:10.1177/1078155215577808.
16. Schwartz RH. Enhancing children's satisfaction with antibiotic therapy: A taste study of several antibiotic suspensions. Current Therapeutic Research. 2000;61(8):570-81. doi:10.1016/s0011-393x(00)80039-9.
17. Thompson A, Reader S, Field E, Shephard A. Open-label taste-testing study to evaluate the acceptability of both strawberry-flavored and orange-flavored amylmetacresol/2,4-dichlorobenzyl alcohol throat lozenges in healthy children. Drugs in R and D. 2013;13(2):101-7.
18. Sjövall J, Fogh A, Huitfeldt B, Karlsson G, Nylén O. Methods for evaluating the taste of paediatric formulations in children: A comparison between the facial hedonic method and the patients' own spontaneous verbal judgement. European Journal of Pediatrics. 1984;141(4):243-7.
19. Reader S, Shaw H, Hails S. taste-testing study in healthy volunteers (children) to investigate children's preference for ibuprofen or placebo suspension. 2006;7(2).
20. Smith CJ, Sammons HM, Fakis A, Conroy S. A prospective study to assess the palatability of analgesic medicines in children. Journal of Advanced Nursing. 2013;69(3):655-63.
21. Angelilli ML, Toscani M, Matsui DM, Rieder MJ. Palatability of oral antibiotics among children in an urban primary care center. Arch Pediatr Adolesc Med. 2000;154(3):267-70.
22. Gooch WM, Philips A, Rhoades R, Rosenberg R, Schaten R, Starobin S. Comparison of the efficacy, safety and acceptability of cefixime and amoxicillin/clavulanate in acute otitis media. The Pediatric Infectious Disease Journal. 1997;16(2):S21-S4.
23. Matsui D, Lim R, Tschen T, Rieder MJ. Assessment of the palatability of β-lactamase-resistant antibiotics in children. Archives of Pediatrics and Adolescent Medicine. 1997;151(6):599-602.
24. Powers JL, Gooch Iii WM, Oddo LP. Comparison of the palatability of the oral suspension of cefdinir vs. amoxicillin/clavulanate potassium, cefprozil and azithromycin in pediatric patients. Pediatric Infectious Disease Journal. 2000;19(12 SUPPL.):S174-S80.
25. Toscani M, Drehobl M, Freed J, Stool S. A multicenter, randomized, comparative assessment in healthy pediatric volunteers of the palatability of oral antibiotics effective in the therapy of otitis media. Current Therapeutic Research - Clinical and Experimental. 2000;61(5):278-85.
26. Ubaka CM, Udeogaranya OP, Ezeugwu N. Palatability of oral paediatric antibiotics commonly prescribed in a Nigerian specialist hospital. European Journal of Hospital Pharmacy: Science and Practice. 2013;20(2):122-4.
27. Dagan R, Shvartzman P, Liss Z. Variation in acceptance of common oral antibiotic suspensions. Pediatr Infect Dis J. 1994;13(8):686-90.
28. Hames H, Seabrook JA, Matsui D, Rieder MJ, Joubert GI. A palatability study of a flavored dexamethasone preparation versus prednisolone liquid in children with asthma exacerbation in a pediatric emergency department. Canadian Journal of Clinical Pharmacology. 2008;15(1):e95-e8.
29. Isa JM, Wong GK, Teraoka SS, Sera MJ, Tsushima MM, Yamamoto LG. Parental pediatric corticosteroid preferences. The American Journal of Emergency Medicine. 2001;19(1):29-31. doi:http://dx.doi.org/10.1053/ajem.2001.20025.
30. Kim MK, Yen K, Redman RL, Nelson TJ, Brandos J, Hennes HM. Vomiting of liquid corticosteroids in children with asthma. Pediatric emergency care. 2006;22(6):397-401. doi:10.1097/01.pec.0000221338.44798.6a.

31. Powers JL. Properties of azithromycin that enhance the potential for compliance in children with upper respiratory tract infections. The Pediatric Infectious Disease Journal. 1996;15(9):30-7.
32. Stafford L, Hope ME, Janney EP, Ailakis JG. Comparison of pediatric steroid mixtures. Australian Journal of Hospital Pharmacy. 1998;28(4):246-9.

Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

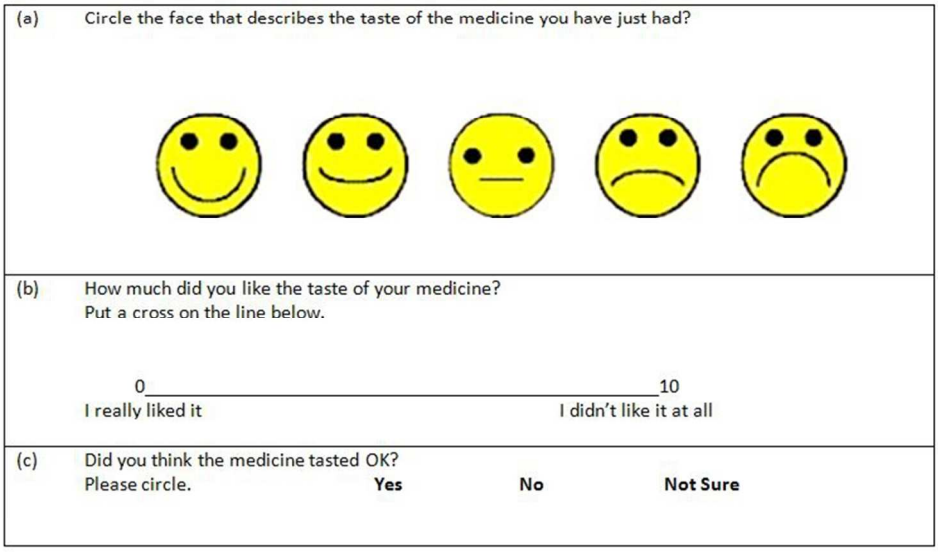


Figure 1. Scales used within PRO tools completed by paediatric participants immediately after administration of their medicine. (a) Hedonic scale; (b) Anchored visual analogue scale (VAS); (c) direct question on taste.

67x53mm (300 x 300 DPI)

Facial expressions observed

Expression	Tick if observed prior to administration	Tick if observed during administration
Eyes squeezed shut or towards shut		
Brow bulge/lower (frown)		
Nose wrinkle		
Pursed lips		

Behaviours observed

Behaviour	Tick if observed prior to administration	Tick if observed during administration	
Child refuses medicine			=unacceptable
Child cries/screams			=unacceptable
Child requires physical restraint			=unacceptable
Child voices resistance			=unacceptable
	Tick if observed immediately after administration		
Child voices disgust			=unacceptable
Child vomits			=unacceptable
Child spits out medicine			=unacceptable
Child cries			=unacceptable

Figure 2. Researcher observation sheet completed by the researcher prior to, during and post medicine administration.

65x41mm (300 x 300 DPI)

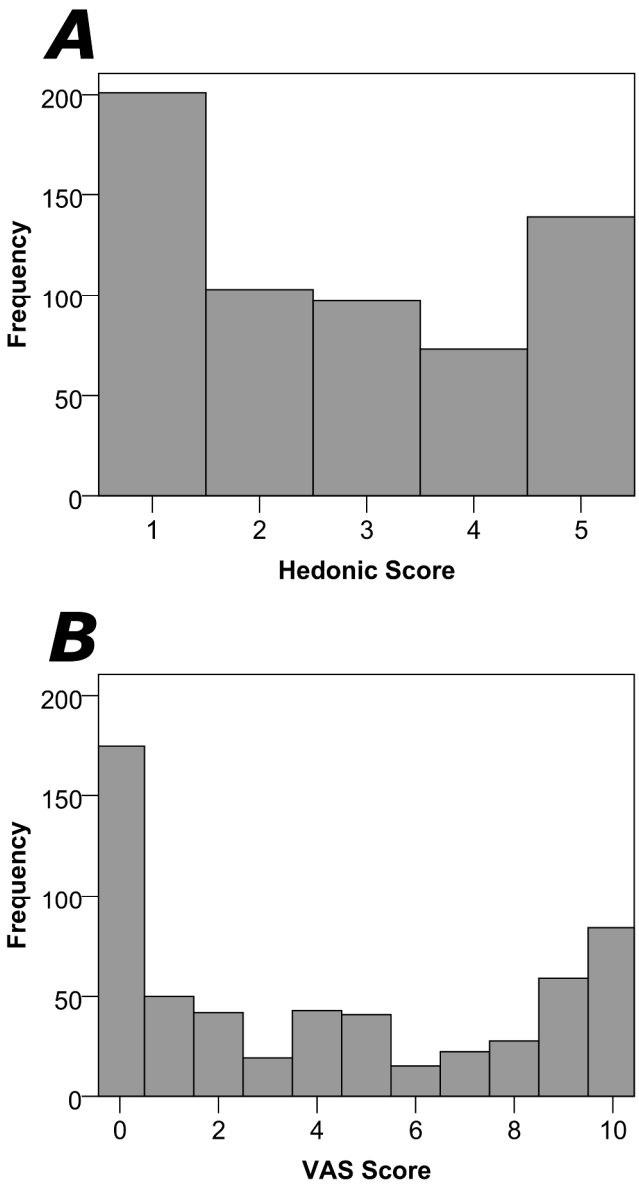
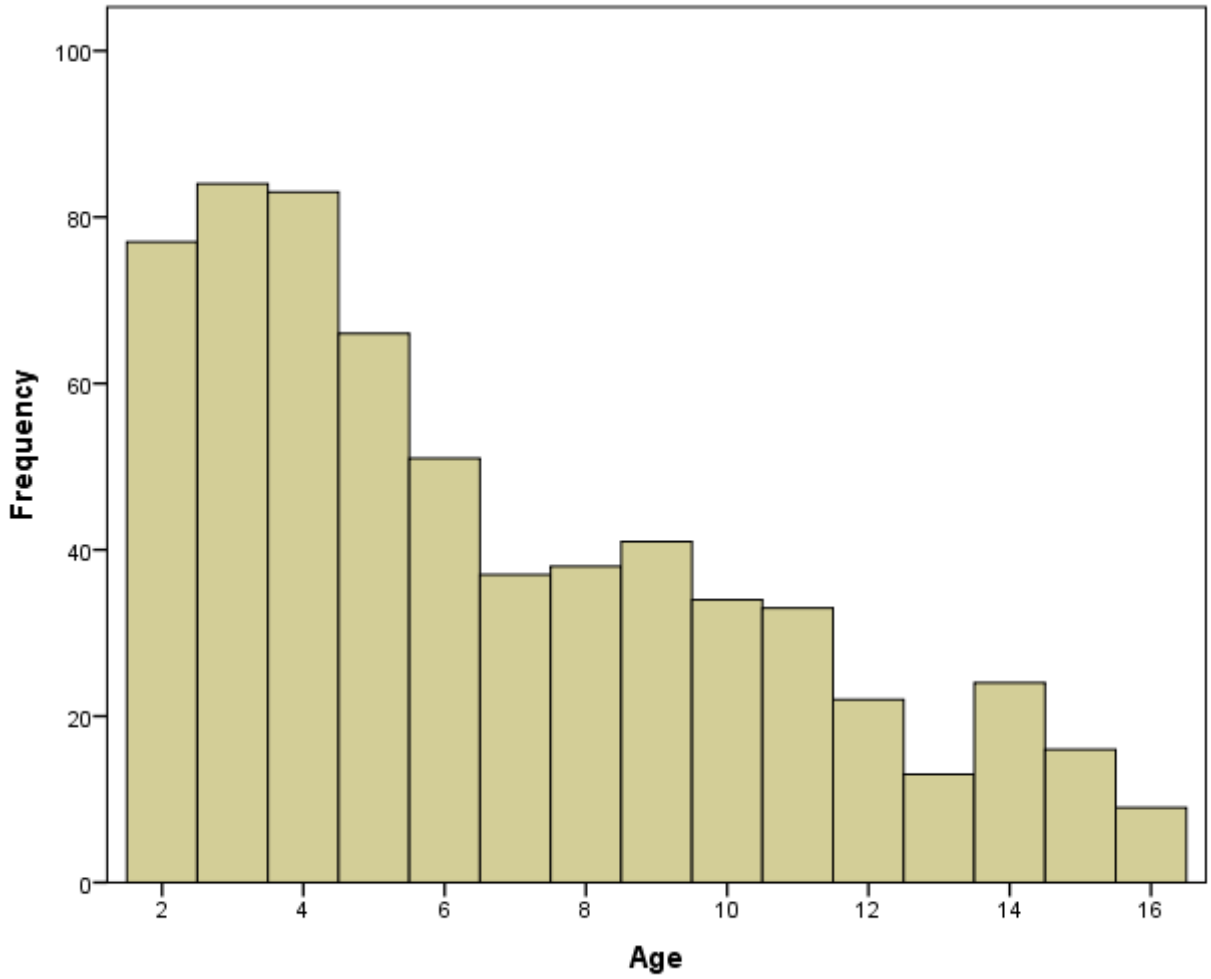
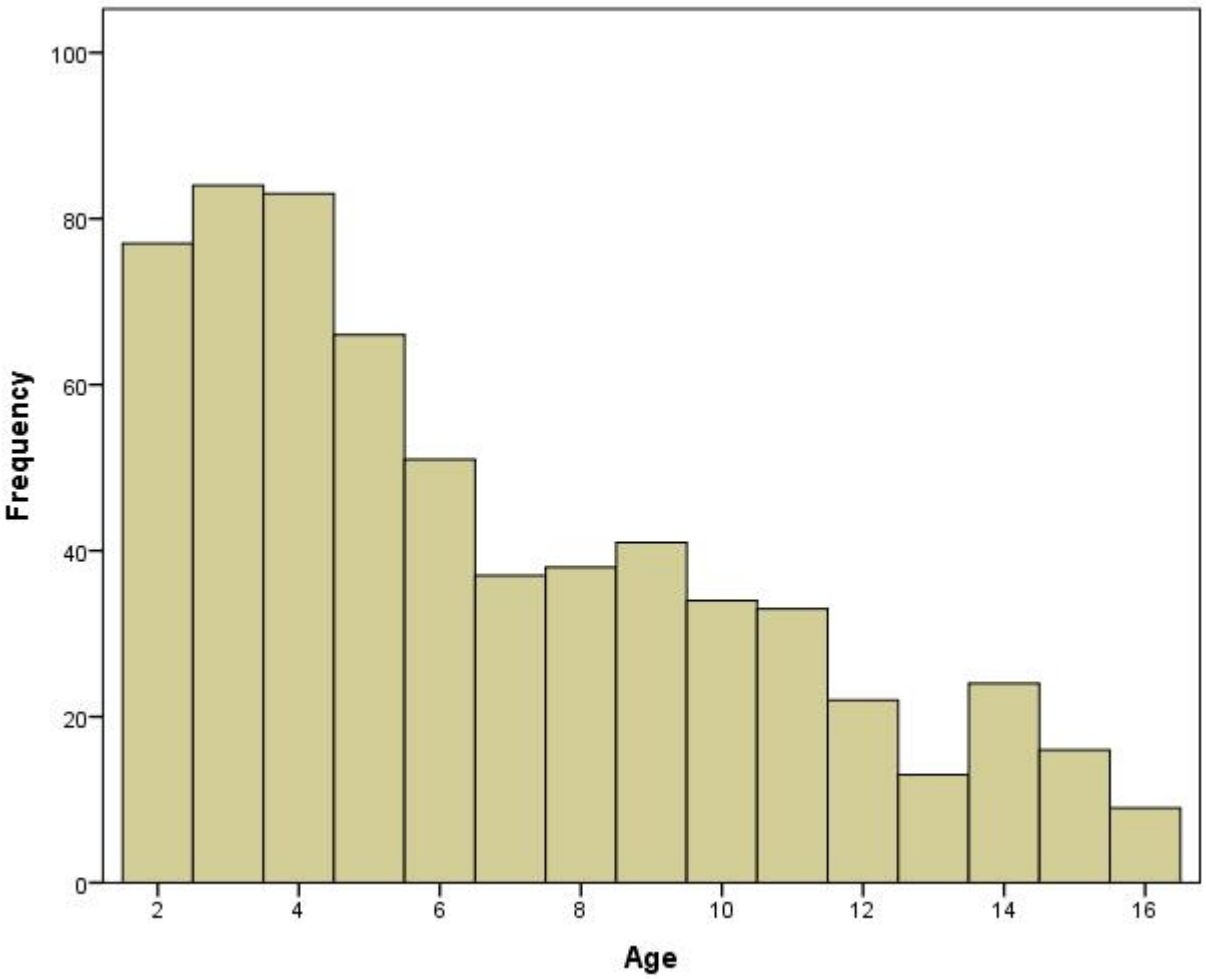


Figure 3. Hedonic and VAS score distribution

168x300mm (300 x 300 DPI)

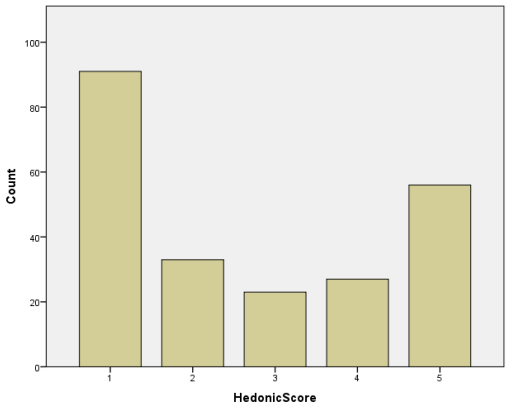
Supplementary Material 1. Distribution of participant age

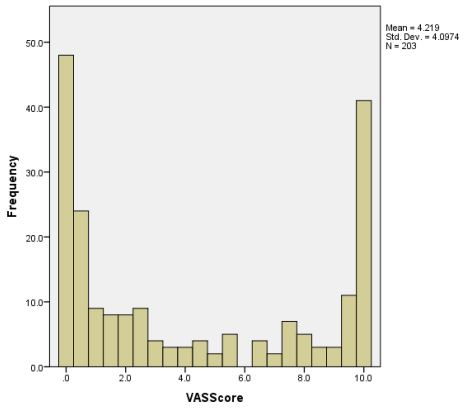




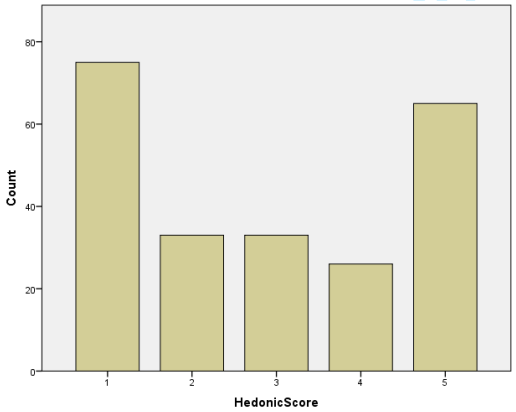
Supplementary Material 2. Age related distribution of responses from patient-reported assessment scales

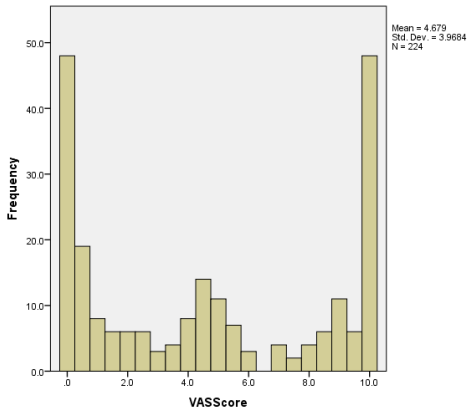
(a) Aged 2-4 years



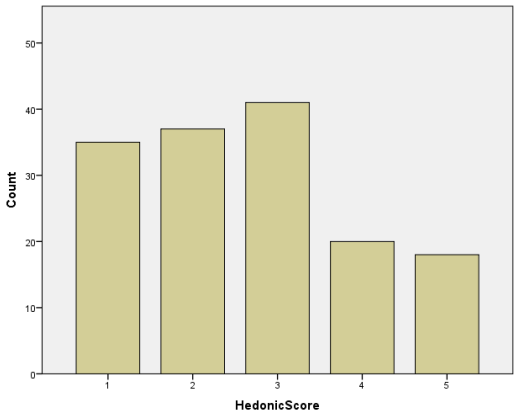


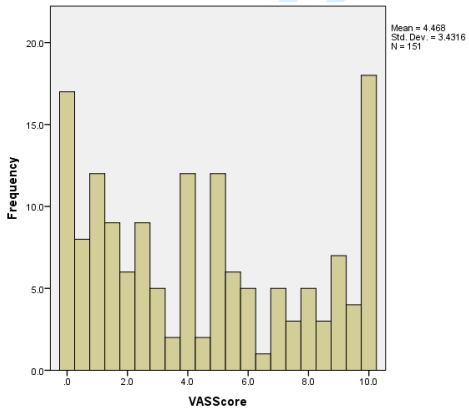
(b) Aged 5-9 years





(c) Aged 10-16 years





Supplementary material 3. Table of taste scores for the six most commonly administered drugs by brand

Drug/Manufacturer	N	Mean Hedonic Score	Mean VAS Score	Tastes OK (% "No")
Amoxicillin				
Athlone Laboratories	21	2.8	4.4	33%
Clamo	2	5.0	0.3	100%
Kent Pharmaceuticals	2	1.0	1.3	0%
Medreich PLC	1	3.0	10.0	100%
PL Holder	4	2.0	1.3	0%
Clarithromycin				
Abbott	11	3.7	7.4	82%
Klaricid	1	5.0	9.9	100%
MA Holder	1	2.0	4.2	0%
Sandoz Ltd.	13	3.4	6.7	54%
Co Amoxiclav				
Medreich PLC	1	3.0	5.0	0%
Rosemont Pharmaceutical Ltd	1	2.0	2.4	0%
Sandoz Ltd.	48	2.5	4.0	30%
Ibuprofen				
Fenpaed	9	2.1	3.4	22%
MA Holder	1	5.0	5.2	0%
McNeil Products	1	5.0	10.0	100%
Pinewood	78	2.0	2.5	10%
Reckitt Benckiser Healthcare	3	2.3	3.7	0%
Paracetamol				
Bristol Laboratories	1	3.0	4.0	100%
MA Holder	38	1.9	2.4	3%
McNeil Products	6	2.3	2.5	17%
Pinewood	113	2.6	4.3	28%
Rosemont Pharmaceutical Ltd	35	2.4	3.0	17%
Prednisolone				
Actavis	12	3.1	4.7	42%
Amdipharm	64	3.6	7.0	63%
Genetic SPA	1	3.0	5.2	0%
Teva UK Ltd	9	4.5	8.6	86%

STROBE Statement—checklist of items that should be included in reports of observational studies

This checklist was used and all items are provided within the manuscript: *Evaluation of patient reported outcome measurements as a reliable tool to measure acceptability of the taste of paediatric medicines*

	Item No	Recommendation	Appears in manuscript (line number)
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	34-35
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Page 2 – Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	115-123
Objectives	3	State specific objectives, including any prespecified hypotheses	139-142
Methods			
Study design	4	Present key elements of study design early in the paper	145-150
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	167-180
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	168-171
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	183-208
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*172-174; 183-208*
Bias	9	Describe any efforts to address potential sources of bias	211-212
Study size	10	Explain how the study size was arrived at	216-218
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings	193-196; 219-232

		were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	220-222; 215-232
		(*b*) Describe any methods used to examine subgroups and interactions	240-241
		(*c*) Explain how missing data were addressed	227-228
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	222-224
		(*e*) Describe any sensitivity analyses	284-290

Continued on next page

Results			**Appears in manuscript**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	250-259
		(b) Give reasons for non-participation at each stage	250-259
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	237-239
		(b) Indicate number of participants with missing data for each variable of interest	250-259
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*250-259*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	261-364
		(*b*) Report category boundaries when continuous variables were categorized	301-313
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	343-364
Discussion			
Key results	18	Summarise key results with reference to study objectives	366-422
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	56-69
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	425-438
Generalisability	21	Discuss the generalisability (external validity) of the study results	433-438
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	90-95

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.